

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 13, 2005, 11:35:21 ; Search time 42 Seconds
(without alignments)
1278.308 Million cell updates/sec

Title: US-09-189-415b-11
Perfect score: 2840
Sequence: 1 MPIGNLGNPNVNSIPPAP.....SNSAVNTSNPPAGSHRFV 558

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: Pir1:*
2: Pir2:*
3: Pir3:*
4: Pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2840	100.0	558	2	A98199	translocated intim
2	2840	100.0	558	2	E86045	probable transloca
3	200	7.0	2232	2	T34434	hypothetical prote
4	195.5	6.9	1229	2	T25697	hypothetical prote
5	187	6.6	1291	2	T13389	hypothetical prote
6	172.5	6.1	1630	2	A53577	ascites sialoglyco
7	172.5	6.1	2271	2	F90073	hypothetical prote
8	167.5	5.9	1192	2	T18611	probable serine/th
9	167	5.9	2468	2	A83412	hypothetical prote
10	166	5.8	796	2	T21460	hypothetical prote
11	164	5.8	971	2	B90835	hypothetical prote
12	164	5.8	973	2	C85693	probable tail fibe
13	164	5.8	1275	2	T33369	probable membrane
14	164	5.8	1770	2	A71517	hypothetical prote
15	163.5	5.8	1246	2	G89287	hypothetical prote
16	163.5	5.8	1802	2	S69703	HKR1 protein precu
17	163	5.7	786	2	T16509	hypothetical prote
18	160	5.6	1122	2	G64887	probable tail fibe
19	160	5.6	1829	2	T24583	hypothetical prote
20	158.5	5.6	1063	2	D86731	hypothetical prote
21	157.5	5.5	3013	2	A80480	probable invasin y
22	156.5	5.5	3570	2	T45025	mucin MUC5B, trach
23	155.5	5.5	3507	2	T34513	hypothetical prote
24	152.5	5.4	1077	2	A44067	serine-rich protei
25	152	5.4	461	2	JN0097	secreted 45K prote
26	152	5.4	997	2	T43523	cut17 protein - fi
27	151.5	5.3	2660	2	E85882	probable invasin z
28	151	5.3	918	2	T02759	hypothetical prote
29	151	5.3	1306	2	S25370	MSB2 protein - yea

30	150	5.3	888	2	T46726	secreted acid phos
31	150	5.3	1026	1	A40315	maternal effect pr
32	150	5.3	1034	2	JC2143	ice nucleation act
33	150	5.3	1460	2	D81675	polymorphic membra
34	149.5	5.3	1772	2	A45532	major merozoite su
35	149.5	5.3	2514	2	F81045	hemagglutinin/hemo
36	148.5	5.2	770	2	T51024	related to C2H2 zi
37	148.5	5.2	1189	2	S56852	hypothetical prote
38	148	5.2	2035	2	A40718	host cell factor C
39	148	5.2	3190	2	T13828	CREB-binding prote
40	147.5	5.2	797	1	VBEX1	glycoprotein X pre
41	147.5	5.2	1547	2	T28657	blackjack protein,
42	147.5	5.2	2055	2	T31110	extracellular matr
43	147.5	5.2	2059	2	D82671	surface protein XF
44	147.5	5.2	3624	2	AD0835	large repetitive p
45	147.5	5.2	4558	2	C82199	RTX toxin RtxA VCI

ALIGNMENTS

RESULT 1

A:98199

translocated intimin receptor Tir [imported] - Escherichia coli (strain O157:H7, substrain C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Jul-2004
C:Accession: A98199
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.; gasawara, N.; Yasunaga, T.; Shiba, S.; Hattori, M.; Shingawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genom
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: A98199
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-558 <HAY>
A:Cross-references: UNIPROT:Q9R396; GB:BA000007; PIDN:BA037984.1; PID:gl3364036; GSPDB:G
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: EC94561

Query Match	100.0.0%;	Score	2840;	DB	2;	Length	558;
Best Local Similarity	100.0.0%;	Pred. No.	3.4e-149;				
Matches	558;	Conservative	0;	Mismatches	0;	Indels	0;
QY	1	MPICNLGNHPNVNNSIPPAPLP	QSOTD GAGRGQLINSTGPLGSRALFTPVNRSMDSGD	60			
DB	1	MPICNLGNHPNVNNSIPPAPLP	QSOTD GAGRGQLINSTGPLGSRALFTPVNRSMDSGD	60			
QY	61	NRASDVPLGPNVPMRLAAASEITLNDGFVLHDHGPLD	TLNRQIGSSVFRVETQEDGKHIA	120			
DB	61	NRASDVPLGPNVPMRLAAASEITLNDGFVLHDHGPLD	TLNRQIGSSVFRVETQEDGKHIA	120			
QY	121	VGQRNGVTSVVLSDQEVARLQSIDPEGKDKFVFTGGRGGAGHAMVT	VASDITEARQIL	180			
DB	121	VGQRNGVTSVVLSDQEVARLQSIDPEGKDKFVFTGGRGGAGHAMVT	VASDITEARQIL	180			
QY	181	ELLEPKGTGESKGAGESKGVGELRESNSGAENTTETQTSTSSLRSDPKLWALGT	VTAT	240			
DB	181	ELLEPKGTGESKGAGESKGVGELRESNSGAENTTETQTSTSSLRSDPKLWALGT	VTAT	240			
QY	241	GLIGLAATGIQVLAALTPEDSPSTTTTDDPAAASATETATRDQLTKEAFONPDNQKNIDE	300				
DB	241	GLIGLAATGIQVLAALTPEDSPSTTTTDDPAAASATETATRDQLTKEAFONPDNQKNIDE	300				
QY	301	LGNAPISVGLKDDVVAIEEQAAKAGEEAKQQA	IENNAQAQKYDEQAQKEELKVSSG	360			
DB	301	LGNAPISVGLKDDVVAIEEQAAKAGEEAKQQA	IENNAQAQKYDEQAQKEELKVSSG	360			
QY	361	AGYGLSGALILGGGIGVAVTAALHRKNQPV	EQTTTTTTTTTTTTTSARTVENKPANNTPAQG	420			
DB	361	AGYGLSGALILGGGIGVAVTAALHRKNQPV	EQTTTTTTTTTTTTTSARTVENKPANNTPAQG	420			

QY 421 NVDTGSEDTMESRRSSMASTSTFFDTSSIGTVQNPYADVKTSLHDSQVPTSNSTSVQ 480
Db |||||||
QY 421 NVDTGSEDTMESRRSSMASTSTFFDTSSIGTVQNPYADVKTSLHDSQVPTSNSTSVQ 480
Db |||||||
QY 481 NMGNTDSVVYSTIQQHPRDITDNGARLLGNPAGIQSTYARLALSGGLRHDGGLTGGSN 540
Db |||||||
QY 481 NMGNTDSVVYSTIQQHPRDITDNGARLLGNPAGIQSTYARLALSGGLRHDGGLTGGSN 540
Db |||||||
QY 541 SAVNTSNPPAPGSHRFV 558
Db |||||||
QY 541 SAVNTSNPPAPGSHRFV 558
Db |||||||
RESULT 2
E86045
probable translocated intimin receptor protein tir [imported] - Escherichia coli (strain
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: E86045
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glaesner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouais, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: E86045
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-558 <STO>
A:Cross-references: UNIPROT:Q9R396; GB:AE005174; NID:g12518449; PIDN:AGS58825.1; GSPDB:G
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: tir
Query Match 100.0%; Score 2840; DB 2; Length 558;
Best Local Similarity 100.0%; Pred. No. 3.4e-149;
Matches 558; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MPIGNLHNPNVNNISIPAPPPLPSQTDGAGRGQLINSTGPGSRALFTPVNSMADSGD 60
Db |||||||
QY 1 MPIGNLHNPNVNNISIPAPPPLPSQTDGAGRGQLINSTGPGSRALFTPVNSMADSGD 60
Db |||||||
QY 61 NRASDVGLPVNPMLAASEITLNDGFVLDHGPDLTLNRQIGSSVFRVETQEDGKHIA 120
Db |||||||
QY 61 NRASDVGLPVNPMLAASEITLNDGFVLDHGPDLTLNRQIGSSVFRVETQEDGKHIA 120
Db |||||||
QY 121 VQORNGVETSVLSDOEVARLOSIDPEGKDFVFTGGRGGAGHAMVTVASDITEARQIL 180
Db |||||||
QY 121 VQORNGVETSVLSDOEVARLOSIDPEGKDFVFTGGRGGAGHAMVTVASDITEARQIL 180
Db |||||||
QY 181 ELLBPKGTGESKGESKGVGLRESNSGAENTTETQTSSTSLRSDPKLWALGTAT 240
Db |||||||
QY 181 ELLBPKGTGESKGESKGVGLRESNSGAENTTETQTSSTSLRSDPKLWALGTAT 240
Db |||||||
QY 241 GLIGLAATGIVOALALTPEPDSPTTDPDAAASATETATRDQLTKEAFQNPQNKVNIDE 300
Db |||||||
QY 241 GLIGLAATGIVOALALTPEPDSPTTDPDAAASATETATRDQLTKEAFQNPQNKVNIDE 300
Db |||||||
QY 301 LGNAIPSGVLKDDVYANIEEQAKAGEAKQOAIENNAQAQKYDEQAKQEELKVSSG 360
Db |||||||
QY 301 LGNAIPSGVLKDDVYANIEEQAKAGEAKQOAIENNAQAQKYDEQAKQEELKVSSG 360
Db |||||||
QY 361 AGYGLSGALILGGGIGVATTAALHRKNQPVETTTTTTTTTTSARTVENKPNNTPAQG 420
Db |||||||
QY 361 AGYGLSGALILGGGIGVATTAALHRKNQPVETTTTTTTTTTSARTVENKPNNTPAQG 420
Db |||||||
QY 421 NVDTGSEDTMESRRSSMASTSTFFDTSSIGTVQNPYADVKTSLHDSQVPTSNSTSVQ 480
Db |||||||
QY 421 NVDTGSEDTMESRRSSMASTSTFFDTSSIGTVQNPYADVKTSLHDSQVPTSNSTSVQ 480
Db |||||||
QY 481 NMGNTDSVVYSTIQQHPRDITDNGARLLGNPAGIQSTYARLALSGGLRHDGGLTGGSN 540
Db |||||||
QY 481 NMGNTDSVVYSTIQQHPRDITDNGARLLGNPAGIQSTYARLALSGGLRHDGGLTGGSN 540
Db |||||||

QY 541 SAVNTSNPPAPGSHRFV 558
Db |||||||
QY 541 SAVNTSNPPAPGSHRFV 558
Db |||||||
RESULT 3
T34434
hypothetical protein K06A9.1a - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T34434
R:Geisler, C.; Gattung, S.
submitted to the EMBL Data Library, December 1996
A:Description: The sequence of C. elegans cosmid K06A9.
A:Reference number: Z21525
A:Accession: T34434
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2232 <GEI>
A:Cross-references: UNIPROT:Q8IFX6; EMBL:U80846; PIDN:AAC70890.1; GSPDB:GN00028; CESP:K06
A:Experimental source: strain Bristol N2; clone K06A9
C:Genetics:
A:Gene: CESP:K06A9.1a
A:Map position: X
A:Introns: 38/1; 75/3; 103/3; 132/2; 158/2; 222/1; 1088/1; 1367/1; 2039/1; 2049/1; 2075/1;
Query Match 7.0%; Score 200; DB 2; Length 2232;
Best Local Similarity 21.2%; Pred. No. 0.0061;
Matches 127; Conservative 80; Mismatches 278; Indels 114; Gaps 21;
QY 9 NPVNNISIPAPPPLPSQTDGAGRGQLINSTGPGSRALFTPVNSM----- 55
Db |||||||
QY 1343 SPSSISPVPTSSPIPTTFASSTSGSTISDVSSVSTSL-APLSSSLSTVPSSSQSFSS 1401
Db |||||||
QY 56 ADSGNRASDVP-----GLPNPMLAASEITLNDGFVLDHGPDLTLNRQIGSSVFR 109
Db |||||||
QY 1402 TSESSKASSPVPSTSTPTNPTGSESTSLSSITSGSTQH---TTMSKASSGSP 1458
Db |||||||
QY 110 VETQEDGKHIAVG--QRNGVETSVLSDOEVARLOSIDPEGKDFVFTGGRGGAGHA--- 164
Db |||||||
QY 1459 STNSQTSVTVMGSSSTSGVSTSGASTQPMQSTSQSSAGSTVASSTASPAASSTAPSS 1518
Db |||||||
QY 165 -----MVTVASDITEARQILELLEPKGTGESKGESKGVGLRESNSGAENTTET 216
Db |||||||
QY 1519 TGTWSSSTSGTVGSTISES-STTASASSQTSVTVMGSSSTSGV-----STSSASSTQ 1571
Db |||||||
QY 217 QTSSTSLRSDPKLWALGTATGLIGLAATGIVOALALTPEPDSPTTDPDAAASATE 276
Db |||||||
QY 1572 QMSTSQGS-----SAGSTVASSTAGLVSTTV-----PSSTGTMGSTSGTVGS 1615
Db |||||||
QY 277 TATRDQLTKEAFQNPQNKVNIDELGNAPSGVLKDDVYANIEEQAKAGEAKQOAIEN 336
Db |||||||
QY 1616 TISESSTTASA-----SSQTSVTVMGSSSTSGV-----STSSASSTQPMQSTSQ 1660
Db |||||||
QY 337 NAOQKKYDEQAKQEELKVSSGAG-YGLSCALITGGGIGVAVTA--ALHRKNQPVQ 393
Db |||||||
QY 1661 GSSAGSTVASSTGTLVSTSTVPSTGTMGSTSGTVGSTISESSTTASASSQTSVTMG 1720
Db |||||||
QY 394 TTTTITTTTGA-----RTVENKPNNTPAQNVDPGSEDTMESRRSSMASTSTFF 446
Db |||||||
QY 1721 SSSTSGVSTSSASSGQPMQSTSQSSAGSTVVSSTA-SPAASSTAPSTGTMSTSS--- 1776
Db |||||||
QY 447 DTSSITGV-----QNPYADVKTSLHDSQVPTSNSTSVQNMGT-----DSVVYSTIOH 495
Db |||||||
QY 1777 -----GTVGTMSSOSTTAASTTSHGTGTVTIGSSSTSNQMSSTQSSGVSTVASSTAGL 1831
Db |||||||
QY 496 PPRDTTNGARLLGNPSAGI-----OSTYARLALSGGLRHDGGLTGGSNSAVNTS 547
Db |||||||
QY 1832 VSTSTVPSSTGTMGSTSGTVGSTISESSTTASASSQTSVTMG---SSSTSGVSTSS 1887
Db |||||||
RESULT 4
T25697

hypothetical protein F16F9.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T25697

R:Fulton, B.

submitted to the EMBL Data Library, August 1996

A:Description: The sequence of C. elegans cosmid F16F9.

A:Reference number: Z20071

A:Accession: T25697

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1229 <FUL>

A:Cross-references: UNIPROT:O94185; EMBL:U67956; PIDN:AA07691.1; GSPDB:GN00028; CESP.F1

A:Experimental source: strain Bristol N2; clone F16F9

C:Genetics:

A:Gene: CESP.F16F9.2

A:Map position: X

A:Introns: 35/1; 361/1; 384/1; 482/2; 517/1; 971/1; 1021/2; 1179/3

Query Match 6.9%; Score 195.5; DB 2; Length 1229;
Best Local Similarity 22.0%; Pred. No. 0.0048;
Matches 90; Conservative 54; Mismatches 155; Indels 111; Gaps 13;

QY 167 TVASDITEARQRIELLEPKGTGSGKAGESKGVGELR-----ESNSGAENTTETOTSTST 222

DB 597 SVSTKSTTKKASTTEPTTDEPTTTTSTTGKATPELSTTSEETTTTELKITTEG 656

QY 223 SSLRSDPKLWALGTVATGLIGLAATGIVQALALTPEDSPPTTDPDAASATETATRDQ 282

DB 657 SITTEP-----TTTAFABASTGII-----TTDEETTTSTTTPETSTKEIVTESA 704

QY 283 LTKEAF-----QNPQNKVNDLGNALPSGVLDKDV-VANIEQAKAGEAKQO 332

DB 705 ITQTSVSVVESSTPQLPRKAIKVNFKHNL--VLKEKRLLEKESTSTGSDSS-- 760

QY 333 ATENNAQAKKYDEQAKQOEKLVSSGAGYGLSGALILGGGIGVAVTAALHRKNQPVQ 392

DB 761 --ETTVVAENIDEVTTTEKEV-----VQTPITTEKSTTQEE 797

QY 393 TTTTTTTTTTARTVENKPNANTPAQGNVDTPGSEDWMSRRSSMASTSTFFDTSSIG 452

DB 798 TTTTTTTTTEKTSKTTTEKPTTSESA-----TTTETTSPTSITSTVDTSSAT 846

QY 453 TVQNPYA--DVKTSLHDSQVPTNS-----NTSVQNMG----- 483

DB 847 TEESSATAETTTTSAETSTTTSAAFTIGESPENTALQSSQKSEENESSAEKPGARR 906

QY 484 -----NTDSVVVSTIQHPPTDNDNGARLLGNPSAGIQST 518

DB 907 DVPKXKHTTVKPAETTSAAASTTTTTEPITTEKSTTLETTP---IEAT 953

RESULT 5

T13389

hypothetical protein 115C2.10 - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster

C>Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004

C:Accession: T13389

R:Salles, C.; Valenti, P.; Darlanitso, A.; Henderson, N.; Campbell, L.; Glover, D.

submitted to the EMBL Data Library, May 1999

A:Description: Sequencing the distal X chromosome of Drosophila melanogaster.

A:Reference number: Z17665

A:Accession: T13389

A>Status: preliminary;

A:Molecule type: DNA

A:Residues: 1-1291 <CAT>

A:Cross-references: UNIPROT:O77261; EMBL:AL031581; NID:e1320978; PID:e1426292; PIDN:CAA2

C:Genetics:

A:Cross-references: FlyBase:FBgn020381

A:Map position: X

A:Introns: 238/3; 1225/1

A>Note: EG:115C2.10

Query Match 6.6%; Score 187; DB 2; Length 1291;
Best Local Similarity 20.4%; Pred. No. 0.015;
Matches 130; Conservative 74; Mismatches 245; Indels 188; Gaps 25;

QY 5 NLGHNPNVNSIPAPPPLPSQTDG---AGGRGQLINSTGPLGSRALFTPVNRNSMADSGDN 61

DB 438 NSTSNSNTNDSTGPGSETSTINGLVASGGAG-----GATGAAMLPTP---SQQSTGCK 488

QY 62 RAS-----DVPGLPVNPMRLA-----ASEITLNDGPEVLHDHGPLOTLNQ 102

DB 489 EATAAVSLLEKKLPNVVVSPLTMKELRQKGMTKYDAEMIMANAAVQQQHHH----- 539

QY 103 IGSSVFRVETQEDGKHIAVGQRNGVETSVLVSDQYARLOSIDPQKOKFVETGGRGGAG 162

DB 540 --QHFFHHHHHHHHHHHGHQASTGAETAATAVQOQMAAQKPG-----VGGTGAAG 588

QY 163 HAMVTVASDITEARQRIELLEPKGTGSGKAGESKGVGELRESNSGAENTTETOTSTST 222

DB 599 NAGATTVSSVA-----AGAGSEVNGRSTSLRKSNRVNS-----TSSSI 627

QY 223 SSLRSDPKLWALGTVATGLIGLAATGIVQALALTPEDSPPTTDPDAASATETATRDQ 282

DB 628 STASADEVI-----APVVAASISLPSKAPVVLMPRCCKPAQMAIAALHQ 670

QY 283 LTKEAFQNPDKVNIDELGNALPSGVLDKDVVANIEEQAKAGEAKQO-----A 333

DB 671 SQORQLRSEKQEKLTLD-----GESSD---TSSEQKK-----EQQDQDHLQKMF 716

QY 334 TENNAQAKKYDEQAKQOEKLVSSGAGYGLSGALILGGGIGVAVTAALHRKNQPVQ 393

DB 717 LAEEPQPEKS--EEKQEQKQKRVTRNSAGRVGL-----VARLATAHNN-----IA 760

QY 394 TTTTTTTTTTARTV---ENKPNANTPAQGNVDTPGSEDWMSRRSSMASTSTFFDTSS 450

DB 761 TTTNSSSSNKAATTITCNVHNNSNRINHNLSRLSVKSRKPAPSEASSIPSTSS 820

QY 451 IGTQV-----NPVADVKTSLHDSQVPTNSNTSVQNMGNTDSVVYSTIOH----- 495

DB 821 ENQQQATRRSCPTPAYKKNLLASFDPPTSTQIQEQLKDESVTYSPVKGRSRAAA 880

QY 496 -----PPRDTTNDNGARL--LGNPSAGIQST-----VARLALSGG 527

DB 881 LAAQSIHCEALGPGFTGTSQKRAQAGFTTSCSTTISNVPELLKTPERRKLTLR 940

QY 528 LRH-----DMGLTGGSNSAVNTSNPPAGPSHR 556

DB 941 MKRSPILDEVIELG--TSLSNGGAGRG---APGSHR 971

RESULT 6

A53577

ascitesialoglycoprotein 1 - rat (fragments)

C:Species: Rattus norvegicus (Norway rat)

C>Date: 12-Apr-1995 #sequence_revision 12-Apr-1995 #text_change 07-Feb-1997

C:Accession: A53577

R:Wu, K.; Fregien, N.; Carraway, K.L.

J. Biol. Chem. 269, 11950-11955, 1994

A:Title: Molecular cloning and sequencing of the mucin subunit of a heterodimeric, bifun

A:Reference number: A53577; PMID:94216302; PMID:8163496

A:Accession: A53577

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1630 <WUA>

A:Cross-references: GB:U06746

C:Keywords: glycoprotein

Query Match 6.1%; Score 172.5; DB 2; Length 1630;

Best Local Similarity 20.0%; Pred. No. 0.13;

Matches 136; Conservative 70; Mismatches 272; Indels 203; Gaps 23;

QY 12 VNNISIPAP-----PLPSQTDGAGGRGQLINSTGPLGSRALFTPVNRNSMADSGDNASDVP 67

DB 28 VNTSTTSAPKTSALPSSQMTSNQVSNPTA-----SSYRMTKNTQASPMVTSSST 82

[illegible]

RESULT 9
A83412
hypothetical protein PA1874 [imported] - Pseudomonas aeruginosa (strain PA01)
C/Species: Pseudomonas aeruginosa
C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C/Accession: A83412
R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A/Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A/Reference number: A82950; MUID:20437337; PMID:10984043
A/Accession: A83412
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-2468 <STO>
A/Cross-references: UNIPROT:Q912M3; GB:AE004613; GB:AE004091; NID:q9947856; PIDN:AA05026

[illegible]

```

RESULT 10
T21460
hypothetical protein ZK945.10 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T21460
E:Wilkinson, J
submitted to the EMBL Data Library, March 1995
A:Reference number: Z19425
A:Accession: T21460
A:Status: preliminary; translated from GE/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-796 <WIL>
A:Cross-references: EMBL:Z48582; PIDN:CAA88469.1; GSPDB:GN00020; CESP:ZK945.10
A:Experimental source: clone F27E5
C:Genetics:
A:Gene: CESP:ZK945.10
A:Map position: 2
A:Introns: 62/3; 124/3; 169/1; 216/1; 262/3; 371/1; 581/3; 608/1

```

```
Query Match      5.8%; Score 166; DB 2; Length 796;
Best Local Similarity 20.3%; Pred. No. 0.11;
Matches 108; Conservative 89; Mismatches 223; Indels 112; Gaps 19;

QY 68 GLPVNPMRLAAASEITLND-----GPEVLHDHGPLEDLNFIQIGSSVFRVETQEDGKHI 119
D 137 GFLNSTWITLNEVNDDEISIAVEAKYECVYDDG-----IDRCDSLWM---LQVGGNEM 189
QY 120 A-VGORGNETSVLSDQYARLQSIDPEGDKFVFTGGRGGAGHAMVTVASDITEARQR 178
D 190 ALLGYREKCESEI--NEYARMCKRPYRSK-----STAI SDSQGV 230
QY 179 ILELLEPKGTGSGKAGESKGVGELRESNGAENTTETQTSTSTSSLRSDPKMLALGTV 238
D 231 YVDGQVLKGVRAKQFSMRTSGSPTLRRMKRDAGDNTCDVTIESTSTSTPTTTTSTTV 290
QY 239 -ATGLIGLAATGIVOALALTPEDSPFTT-----DPDAASATETATRDQLTKEAFQ 289
D 291 TSTTTVPTSTVTVMSTSTSTPTSTTSTSTSTSTSTSTSTSTSTSTSTSTSTSTST 350
QY 290 NPDNQKVIDELGNAIPSGVLKD--DVVANIEEQAKAAG-BEAKQQAENNAQAOKKYDE 346
D 351 SPSS-----TTLSTSIPTTTEITSTISLSDPNAICSYLDTTTTSTTTMLTSTTTE 405
QY 347 QOAKRQELKVSSGAGYGLSGALILGGGIGVAVTAALHRKNQPVQTTTTTTTTTTT--S 404
D 406 EPSTSTTTTEVTS-----TSTVTTEPTTTTLTSTASTSTEPS 445
QY 405 ARTVENKPA-----NNTPAQGNVDTPGSEDTMESRRSSMASTSTFTFDTSIGTVQ 455
D 446 TSTVTSTPSTPVTSTVSSSSSTTVTPTSTESTSTSPSTSTVTTTAPSTSTGSPSS 505
QY 456 NPYADVKTSLHDSQVPTNSNTSVQNMGNVTSVYTIQHPRDRTDNGALGNP---- 511
D 506 S-----SSTPSTASSVSTSTASTQSTSTSTSTSTSTSTSTSTSTSTSTSTST 555
QY 512 -----SAGIQSTYARLALSGRLHDMGGLTGSGNSAVNTSNPPAPGSHRFV 558
D 556 VEKATTFPYDSTSVNLNLSGL-----GIIGVQTSIECTS-----PTSSNVV 597

RESULT 11
B90835
probable tail fiber protein [imported] - Escherichia coli (strain RIM
C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: B90835
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: B90835
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-971 <HAY>
A:Cross-references: UNIPROT:Q8XDQ4; GB:BA000007; PIDN:BA035073.1; PID:g13361114; GSPDB:C
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: Ecs1650

Query Match      5.8%; Score 164; DB 2; Length 971;
Best Local Similarity 21.2%; Pred. No. 0.19;
Matches 95; Conservative 66; Mismatches 198; Indels 90; Gaps 15;

QY 79 SEITLNDGFEVLH-----DHGPLEDLNFIQIGSSVFRVETQEDGKHIQVQ----- 123
D 59 SVILLVEGPPSHAGTTIYEDSQP-GTLNDFLGAM-----TEDDVRPEARLRRFELMVEE 112
QY 124 --RNGVETSVVLSDOEYARLQSIDPEGDKFVFTGGRGGAGHAMVTVASDITEARQLLE 181
D 113 VARN---ASAVAQNTAAAKKSASDAS-----TSAREAAATHA---TDAADSARA----- 154

RESULT 12
C85693
probable membrane protein of prophage CP-933X Z1918 [imported] - Escherichia coli (strain
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: C85693
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glaesner, J.D.; Rose, D.J.; Mayhew,
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouais, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: C85693
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-973 <STO>
A:Cross-references: UNIPROT:Q8XDQ4; GB:AE005174; NID:g12514847; PIDN:AAG56007.1; GSPDB:G
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: Z1918

Query Match      5.8%; Score 164; DB 2; Length 973;
Best Local Similarity 21.2%; Pred. No. 0.19;
Matches 95; Conservative 66; Mismatches 198; Indels 90; Gaps 15;

QY 79 SEITLNDGFEVLH-----DHGPLEDLNFIQIGSSVFRVETQEDGKHIQVQ----- 123
D 61 SVILLVEGPPSHAGTTIYEDSQP-GTLNDFLGAM-----TEDDVRPEARLRRFELMVEE 114
QY 124 --RNGVETSVVLSDOEYARLQSIDPEGDKFVFTGGRGGAGHAMVTVASDITEARQLLE 181
D 115 VARN---ASAVAQNTAAAKKSASDAS-----TSAREAAATHA---TDAADSARA----- 156
QY 182 LLEPKGTGSGKAGESKGVGELRESNGAENTTETQTSTSTSSLRSDPKMLALGTVATG 241
D 157 -----ASTSAGCAASSAQSSASSAGTASTKATEASKSAAAAESSK-----SAAAT 201
QY 242 LIGLAATGIVOALALTPEDSPFTTDPDAASATETATRDQLTKEAFONPNQKNIDEL 301
D 202 SAGAAKTSETNVAVSQQAATSTASTATTKAGEAASSARDASASKEAAKSSSTSAAS----- 257
QY 302 GNAIPSGVLKDDVVVANIEEQAKAAGEEAK-QQAATENNAQAOKKYDEQOAKRQEELKVSSG 360
D 258 -----SASSAASSATAAGNSAKAAKTSETNAKSETAQAASASAAAGSK--TA 303
QY 361 AGYGLSGALILGGGIGVAVTAALHRKNQPVQTTTTTTTTTTTTTTTTSARTVENKPNTPAQ- 419
D 304 AALSASAASTSAQOASASATAA-----GKSABSAASSASTATTKAGEATEQAASAASSASA 359
```

```
QY 182 LLEPKGTGSGKAGESKGVGELRESNGAENTTETQTSTSTSSLRSDPKMLALGTVATG 241
D 155 -----ASTSAGCAASSAQSSASSAGTASTKATEASKSAAAAESSK-----SAAAT 199
QY 242 LIGLAATGIVOALALTPEDSPFTTDPDAASATETATRDQLTKEAFONPNQKNIDEL 301
D 200 SAGAAKTSETNVAVSQQAATSTASTATTKAGEAASSARDASASKEAAKSSSTSAAS----- 255
QY 302 GNAIPSGVLKDDVVVANIEEQAKAAGEEAK-QQAATENNAQAOKKYDEQOAKRQEELKVSSG 360
D 256 -----SASSAASSATAAGNSAKAAKTSETNAKSETAQAASASAAAGSK--TA 301
QY 361 AGYGLSGALILGGGIGVAVTAALHRKNQPVQTTTTTTTTTTTTTTTTSARTVENKPNTPAQ- 419
D 302 AALSASAASTSAQOASASATAA-----GKSABSAASSASTATTKAGEATEQAASAASSASA 357
QY 420 ---GNVDTPGSEDTMESRRSSMASTSTFTFDTSIGTVONPYADVKTSLHDSQVPTSN 476
D 358 AKTSETNAKSETSAESSKTAASASSASAASSASASASASASASASASASASASASASAT 417
QY 477 TSVQNMGNVTSVYTIQHPRDRTDNGA 505
D 418 A-----TEAAGSATAAAQSKSTAESAA 439

RESULT 12
C85693
probable membrane protein of prophage CP-933X Z1918 [imported] - Escherichia coli (strain
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: C85693
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glaesner, J.D.; Rose, D.J.; Mayhew,
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouais, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: C85693
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-973 <STO>
A:Cross-references: UNIPROT:Q8XDQ4; GB:AE005174; NID:g12514847; PIDN:AAG56007.1; GSPDB:G
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: Z1918

Query Match      5.8%; Score 164; DB 2; Length 973;
Best Local Similarity 21.2%; Pred. No. 0.19;
Matches 95; Conservative 66; Mismatches 198; Indels 90; Gaps 15;

QY 79 SEITLNDGFEVLH-----DHGPLEDLNFIQIGSSVFRVETQEDGKHIQVQ----- 123
D 61 SVILLVEGPPSHAGTTIYEDSQP-GTLNDFLGAM-----TEDDVRPEARLRRFELMVEE 114
QY 124 --RNGVETSVVLSDOEYARLQSIDPEGDKFVFTGGRGGAGHAMVTVASDITEARQLLE 181
D 115 VARN---ASAVAQNTAAAKKSASDAS-----TSAREAAATHA---TDAADSARA----- 156
QY 182 LLEPKGTGSGKAGESKGVGELRESNGAENTTETQTSTSTSSLRSDPKMLALGTVATG 241
D 157 -----ASTSAGCAASSAQSSASSAGTASTKATEASKSAAAAESSK-----SAAAT 201
QY 242 LIGLAATGIVOALALTPEDSPFTTDPDAASATETATRDQLTKEAFONPNQKNIDEL 301
D 202 SAGAAKTSETNVAVSQQAATSTASTATTKAGEAASSARDASASKEAAKSSSTSAAS----- 257
QY 302 GNAIPSGVLKDDVVVANIEEQAKAAGEEAK-QQAATENNAQAOKKYDEQOAKRQEELKVSSG 360
D 258 -----SASSAASSATAAGNSAKAAKTSETNAKSETAQAASASAAAGSK--TA 303
QY 361 AGYGLSGALILGGGIGVAVTAALHRKNQPVQTTTTTTTTTTTTTTTTSARTVENKPNTPAQ- 419
D 304 AALSASAASTSAQOASASATAA-----GKSABSAASSASTATTKAGEATEQAASAASSASA 359
```

A;Title: Genome sequence of the nematode *C. elegans*: a platform for investigating biology

A>Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_elegans/
 A>Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
 A:Accession: G89287
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1246 <STO>
 A:Cross-references: GB:chr_V; PIDN:CAB09532.1; PID:G3878100; GSPDB:GN00023; CESP:H39E23.
 C:Genetics:
 A:Gene: H39E23.1
 A:Map position: 5

Query Match	5.8%;	Score 163.5;	DB 2;	Length 1246;
Best Local Similarity	19.0%;	Pred. No. 0.28;		
Matches 126;	Conservative 76;	Mismatches 242;	Indels 219;	Gaps 22;

QY	36	INSTGTLGSRALFTPVRSNMDSGDNRASDVGLPNPMLAAASEITLNDGFEVLHDHGP	95
Db	395	INVSSSLGQHPAGVITREHVT-----SSASGSSASPSRYSRSATAT-GASITAGSAL	447
QY	96	LDTLNRQ-----IGSSVFRVETQEDGKHIAVGQNGVETSVVLSDOEYARLQSID	145
Db	448	ASANAQKHQSSAAPSSGSSSSRRSSQNDAATAAG-----GTVMNS-----	490
QY	146	PEGKDKFVTGGRGAGHAMVTVASDITEARQRIELLEPKGTGSKAGESKGVGEURE	205
Db	491	-----GTRHGGVQMQAQT-SRQATISLLQPPSKPSSNTTQIAQIPPLFN	535
QY	206	SNSGAENTTETQTSSTSLR--SDPKLWLALGT-----VATGLIG-----	244
Db	536	RNSTA--TSSAAQPSGTGTTTRKIADPKGRIPLNSTAVQGHRTATGAVAAANNGGIPSHRDH	594
QY	245	-----LAATGIVQALALTPEDSPPTTDPDAAASATETATRDOLTKEAFO-----	289
Db	595	AQQQYMNQLTSSWTMSKLINTPAAGGTAATSSSSSSATSTA---PQKSGSQISHAP	651
QY	290	-----NPDNQKNIDELGNAIPS-----GVLKDDVYANI---EQAKAAGEEA	329
Db	652	TEPVIREDDDENNSENQNGNVLIGGVGQTSQPAQVPTEDATSSSKKEQQQKASSTP	711
QY	330	KQ-----QAIENNAQAQKQDEQAKQOEELKVSAGYGLSGALLILGGGIGVAVTAAL	383
Db	712	KESNPIVWQNLHLNLLKSLDSSAATSYETPRRPGIAG-----	750
QY	384	HRKNQFVEQTTTTTTTTTTTSARTVENKFNNTP-----AQ	419
Db	751	-RRSEPSAATPRRHQTWVDARHLQTPDTPDPYHFEDTILDRQMRALYVSTASSRMT	809
QY	420	GNVDTPGSEDTWESR-----RSMASSTSTFTDTSIGTV-----	454
Db	810	GVLPPTPTSNSTSSSFIVEPLTHVAAASPDITTTTPTKSTVTSYPFRRTPSFRMLIVL	869
QY	455	-----QNPYADVKTSLHDSQVTSNSNTSVQNMGNNTDSVVYSTIOHPERD	499
Db	870	LICDNLRLWPMIHQSP-----SMPPSQMTAMESLKLSESGTGFTGTPATGGPPQRA	923
QY	500	TTDNGARLLGNPSAGIQSTYARLALSGGRHDMGGL---TGGS-----NSAVNTSN	547
Db	924	TSQQMSRSRATTNSANNMGSSGGGAAAAAASATNQLSGAPSTGASSQQYHPKAPSSSSSST	983
QY	548	NPP	550
Db	984	NPP	986

Search completed: May 13, 2005, 11:48:24
 Job time : 46 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 13, 2005, 11:49:17 ; Search time 178 Seconds
(without alignments)
1605.282 Million cell updates/sec

Title: US-09-189-415B-11

Perfect score: 558

Sequence: 1 MPIGNLGHNPVNSIPPAP.....SNSAVNTSNPPAPGSHRFV 558

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1612378 seqs, 512079187 residues

Word size : 8

Total number of hits satisfying chosen parameters: 1770

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : UniProt_03.*

1: uniprot_prot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	558	100.0	558	2	Q7DB77	Q7db77 escherichia
2	558	100.0	558	2	Q9R396	Q9r396 escherichia
3	340	60.9	558	2	O85506	O85506 escherichia
4	28	5.0	547	2	Q7BHL5	Q7bhl5 citrobacter
5	28	5.0	547	2	Q9ET11	Q9et11 citrobacter
6	28	5.0	547	2	Q9WKK1	Q9wkk1 escherichia
7	22	3.9	538	2	O85508	O85508 escherichia
8	22	3.9	538	2	Q47014	Q47014 escherichia
9	22	3.9	538	2	Q47016	Q47016 escherichia
10	20	3.6	367	2	Q79C12	Q79c12 escherichia
11	14	2.5	549	2	O50190	O50190 escherichia
12	14	2.5	550	2	O52147	O52147 escherichia
13	14	2.5	551	2	O68258	O68258 escherichia
14	14	2.5	552	2	Q9KWH9	Q9kwn9 escherichia
15	13	2.3	163	2	Q6GED0	Q6ged0 staphylococ
16	13	2.3	166	2	Q7A060	Q7a060 staphylococ
17	13	2.3	166	2	Q7A2K7	Q7a2k7 staphylococ
18	13	2.3	166	2	Q7A418	Q7a418 staphylococ
19	13	2.3	166	2	Q6G718	Q6g718 staphylococ
20	13	2.3	271	2	Q94491	Q94491 dictyosteli
21	13	2.3	498	2	O86J21	O86j21 dictyosteli
22	12	2.2	56	2	Q8T1V8	Q8tlv8 dictyosteli
23	12	2.2	89	2	Q9NIP9	Q9nip9 trypanosoma
24	12	2.2	106	2	Q6WB03	Q6wb03 trypanosoma
25	12	2.2	106	2	Q6WB04	Q6wb04 trypanosoma
26	12	2.2	106	2	Q6WB05	Q6wb05 trypanosoma
27	12	2.2	107	2	Q6WB01	Q6wb01 trypanosoma
28	12	2.2	107	2	Q6WB08	Q6wb08 trypanosoma
29	12	2.2	107	2	Q9NIQ1	Q9niq1 trypanosoma
30	12	2.2	108	2	Q6WB02	Q6wb02 trypanosoma
31	12	2.2	119	2	Q9N6G3	Q9n6g3 trypanosoma

32	12	2.2	120	2	Q9NIQ0	Q9niq0 trypanosoma
33	12	2.2	130	2	Q6WB06	Q6wb06 trypanosoma
34	12	2.2	132	2	Q9NIQ3	Q9niq3 trypanosoma
35	12	2.2	139	2	Q9C7Q5	Q9c7q5 arabidopsis
36	12	2.2	150	2	Q8LEL8	Q8lel8 arabidopsis
37	12	2.2	200	2	Q9XX76	Q9xx76 caenorhabdi
38	12	2.2	291	2	Q94467	Q94467 dictyosteli
39	12	2.2	336	1	RT09 CANAL	RT09 canal
40	12	2.2	337	2	Q86K30	Q86k30 dictyosteli
41	12	2.2	356	2	Q7PZ21	Q7pz21 anopheles g
42	12	2.2	364	2	Q9W2Z7	Q9w2z7 drosophila
43	12	2.2	402	2	O44626	O44626 caenorhabdi
44	12	2.2	403	2	Q9XJ33	Q9xj33 cyanidium c
45	12	2.2	444	2	Q7S050	Q7s050 neurospora
46	12	2.2	445	2	Q8K3T6	Q8k3t6 mus musculu
47	12	2.2	458	2	Q9NA83	Q9na83 caenorhabdi
48	12	2.2	486	2	Q7S0Z3	Q7s0z3 neurospora
49	12	2.2	524	1	HME2 SCHMA	HME2 schistosoma
50	12	2.2	540	2	Q75JM6	Q75jm6 dictyosteli
51	12	2.2	569	2	Q83151	Q83151 murid herpe
52	12	2.2	569	2	Q83183	Q83183 murid herpe
53	12	2.2	646	2	Q8MNM4	Q8mnk4 dictyosteli
54	12	2.2	648	2	Q86A81	Q86a81 d similar t
55	12	2.2	649	2	Q8MUF9	Q8muf9 dictyosteli
56	12	2.2	659	2	Q94A11	Q94a11 arabidopsis
57	12	2.2	674	2	Q7YZH9	Q7yzh9 monosiga br
58	12	2.2	693	2	Q96XH6	Q96xh6 sulfolobus
59	12	2.2	702	2	Q9V615	Q9v615 drosophila
60	12	2.2	711	2	Q7VYX9	Q7vyx9 cryptospori
61	12	2.2	717	2	Q8U1H5	Q8u1h5 pyrococcus
62	12	2.2	718	2	Q9C6C5	Q9c6c5 arabidopsis
63	12	2.2	722	2	Q9C700	Q9c700 arabidopsis
64	12	2.2	725	2	Q8T124	Q8t124 dictyosteli
65	12	2.2	735	2	Q86LA3	Q86la3 dictyosteli
66	12	2.2	766	2	Q6G7J0	Q6g7j0 drosophila
67	12	2.2	776	2	Q7Q5E1	Q7q5e1 anopheles g
68	12	2.2	798	2	Q86K66	Q86k66 dictyosteli
69	12	2.2	802	1	XYND_RUMFL	XYND_rumfl
70	12	2.2	802	2	Q9S310	Q9s310 ruminococcu
71	12	2.2	832	2	Q86H63	Q86h63 dictyosteli
72	12	2.2	855	2	Q869V0	Q869v0 dictyosteli
73	12	2.2	860	2	Q23916	Q23916 dictyosteli
74	12	2.2	872	2	Q26257	Q26257 dictyosteli
75	12	2.2	986	2	Q9N9W3	Q9n9w3 haemochus
76	12	2.2	1026	1	GLT_DROME	GLT_drome
77	12	2.2	1026	2	Q9VQZ3	Q9vgz3 drosophila
78	12	2.2	1123	2	Q8I811	Q8i811 dictyosteli
79	12	2.2	1247	2	Q9U9T1	Q9u9t1 dictyosteli
80	12	2.2	1247	2	Q9Y0T2	Q9y0t2 dictyosteli
81	12	2.2	1256	2	Q875X6	Q875x6 oikopleura
82	12	2.2	1258	2	Q86KB3	Q86kb3 dictyosteli
83	12	2.2	1297	2	O8SSS5	O8ssss5 dictyosteli
84	12	2.2	1313	2	Q86IU8	Q86iu8 dictyosteli
85	12	2.2	1326	2	Q86I09	Q86i09 dictyosteli
86	12	2.2	1387	2	Q86H9K1	Q86h9k1 dictyosteli
87	12	2.2	1438	2	Q86HL0	Q86hl0 dictyosteli
88	12	2.2	1461	2	Q86AZ9	Q86az9 dictyosteli
89	12	2.2	1565	2	Q86K18	Q86k18 dictyosteli
90	12	2.2	1695	2	Q9NK53	Q9nk53 drosophila
91	12	2.2	1695	2	Q9U9R5	Q9u9r5 drosophila
92	12	2.2	1704	1	VILD_DICDI	VILD_dicdi
93	12	2.2	1711	2	Q9U9R4	Q9u9r4 drosophila
94	12	2.2	1711	2	Q9VJL0	Q9vj10 drosophila
95	12	2.2	1838	2	Q86IV6	Q86iv6 dictyosteli
96	12	2.2	1845	2	Q86JH8	Q86jh8 dictyosteli
97	12	2.2	1909	2	Q86IF3	Q86if3 dictyosteli
98	12	2.2	1929	2	Q8IHL2	Q8ihl2 dictyosteli
99	12	2.2	2159	2	Q86IU5	Q86iu5 dictyosteli
100	12	2.2	2454	2	Q86B02	Q86b02 dictyosteli
101	12	2.2	3432	2	Q95PH9	Q95ph9 dictyosteli
102	12	2.2	3712	1	LMA_DROME	LMA_drome
103	12	2.2	3712	2	Q9VFW0	Q9vfw0 drosophila
104	12	2.2	3770	2	Q869R6	Q869r6 dictyosteli

105	12	2.2	4377	1	ANK3_HUMAN	Q12955 homo sapien	178	11	2.0	165	2	Q94671	Q94671 plasmodium
106	12	2.2	4493	2	Q861H5	Q861h5 dictyosteli	179	11	2.0	166	2	Q7PGC6	Q7pgc6 anopheles g
107	11	2.0	56	2	Q01601	Q01601 pneumocysti	180	11	2.0	172	2	Q9vzb8	Q9vzb8 drosophila
108	11	2.0	56	2	Q861E6	Q861e6 dictyosteli	181	11	2.0	179	2	Q9XZ40	Q9xz40 plasmodium
109	11	2.0	58	2	Q6TUI3	Q6tui3 rattus norv	182	11	2.0	183	1	AACL_DICDI	Pl4195 dictyosteli
110	11	2.0	64	2	Q9N114	Q9n114 plasmodium	183	11	2.0	183	2	Q6UA37	Q6ua97 plasmodium
111	11	2.0	67	2	Q95UY4	Q95uy4 plasmodium	184	11	2.0	183	2	Q9U0B5	Q9u0b5 plasmodium
112	11	2.0	67	2	Q95UY6	Q95uy6 plasmodium	185	11	2.0	183	2	Q9U0B6	Q9u0b6 plasmodium
113	11	2.0	67	2	Q86JN9	Q86jn9 dictyosteli	186	11	2.0	184	2	Q9TY34	Q9ty34 plasmodium
114	11	2.0	71	2	Q962U5	Q962u5 plasmodium	187	11	2.0	184	2	Q9U0B8	Q9u0b8 plasmodium
115	11	2.0	71	2	Q9N1O3	Q9n1o3 plasmodium	188	11	2.0	184	2	Q9U0B9	Q9u0b9 plasmodium
116	11	2.0	71	2	Q6R5F0	Q6r5f0 mus musculu	189	11	2.0	191	2	Q9U0C1	Q9u0c1 plasmodium
117	11	2.0	72	2	Q9W3Q9	Q9w3q9 drosophila	190	11	2.0	195	2	Q93948	Q93948 candida alb
118	11	2.0	74	2	Q61023	Q61023 trypanosoma	191	11	2.0	196	2	Q25947	Q25947 plasmodium
119	11	2.0	74	2	Q9N121	Q9n121 plasmodium	192	11	2.0	196	2	Q25951	Q25951 plasmodium
120	11	2.0	78	2	Q962U6	Q962u6 plasmodium	193	11	2.0	202	2	Q01615	Q01615 pneumocysti
121	11	2.0	86	2	Q9TVF2	Q9tvf2 trypanosoma	194	11	2.0	202	2	Q25952	Q25952 plasmodium
122	11	2.0	86	2	Q61058	Q61058 trypanosoma	195	11	2.0	203	2	Q15806	Q15806 plasmodium
123	11	2.0	98	2	Q9N115	Q9n115 plasmodium	196	11	2.0	203	2	Q9U0C0	Q9u0c0 plasmodium
124	11	2.0	102	2	Q61033	Q61033 trypanosoma	197	11	2.0	203	2	Q9W2W1	Q9w2w1 drosophila
125	11	2.0	107	2	Q61050	Q61050 trypanosoma	198	11	2.0	203	2	Q9LVA3	Q9lva3 arabidopsis
126	11	2.0	108	2	Q9XWNO	Q9xwn0 caenorhabdi	199	11	2.0	204	2	Q8H7E2	Q8h7e2 arabidopsis
127	11	2.0	109	2	Q01619	Q01619 pneumocysti	200	11	2.0	205	2	Q15777	Q15777 trypanosoma
128	11	2.0	109	2	Q6PN90	Q6pn90 plasmodium	201	11	2.0	205	2	Q15911	Q15911 dictyosteli
129	11	2.0	110	2	Q9N116	Q9n116 plasmodium	202	11	2.0	206	2	Q9WFZ9	Q9wzfz9 amsacta alb
130	11	2.0	110	2	Q9N117	Q9n117 plasmodium	203	11	2.0	207	2	Q25701	Q25701 plasmodium
131	11	2.0	110	2	Q9N118	Q9n118 plasmodium	204	11	2.0	208	2	Q25949	Q25949 plasmodium
132	11	2.0	111	2	Q9N126	Q9n126 plasmodium	205	11	2.0	209	2	Q61055	Q61055 trypanosoma
133	11	2.0	112	2	Q6PPZ9	Q6ppz9 plasmodium	206	11	2.0	210	2	Q9Y0Z5	Q9y0z5 trypanosoma
134	11	2.0	113	2	Q9BJN8	Q9bjn8 plasmodium	207	11	2.0	216	2	Q962W6	Q962w6 trypanosoma
135	11	2.0	115	2	Q61046	Q61046 trypanosoma	208	11	2.0	217	1	SGS3_DROSI	P13729 drosophila
136	11	2.0	115	2	Q9B7Q9	Q9bjq9 plasmodium	209	11	2.0	220	2	Q9U0E3	Q9u0e3 plasmodium
137	11	2.0	116	2	Q9B7Q2	Q9bjq2 plasmodium	210	11	2.0	229	2	Q9VIA7	Q9via7 drosophila
138	11	2.0	118	2	Q9NIP8	Q9nip8 trypanosoma	211	11	2.0	236	2	Q7SDS9	Q7sds9 neurospora
139	11	2.0	119	2	Q61034	Q61034 trypanosoma	212	11	2.0	239	2	Q9W2R5	Q9w2r5 drosophila
140	11	2.0	119	2	Q9BJP7	Q9bjp7 plasmodium	213	11	2.0	242	2	Q9VDN0	Q9vdn0 drosophila
141	11	2.0	120	2	Q9GQY0	Q9gqy0 plasmodium	214	11	2.0	244	2	Q9U2Q6	Q9u2q6 caenorhabdi
142	11	2.0	121	2	Q6WAZ9	Q6waz9 trypanosoma	215	11	2.0	245	2	Q9XWP2	Q9xwp2 caenorhabdi
143	11	2.0	122	2	Q15774	Q15774 trypanosoma	216	11	2.0	246	2	Q76P06	Q76p06 dictyosteli
144	11	2.0	122	2	Q9S7P8	Q9s7p8 arabidopsis	217	11	2.0	259	2	Q86IM4	Q86im4 dictyosteli
145	11	2.0	123	2	Q61027	Q61027 trypanosoma	218	11	2.0	260	2	Q8IT83	Q8it83 plasmodium
146	11	2.0	123	2	Q90602	Q90602 trypanosoma	219	11	2.0	264	2	Q96920	Q96920 plasmodium
147	11	2.0	124	2	Q6PQ00	Q6pq00 plasmodium	220	11	2.0	274	1	MSA2_PLAF6	P50497 plasmodium
148	11	2.0	125	2	Q61025	Q61025 trypanosoma	221	11	2.0	274	2	Q7SC14	Q7sc14 neurospora
149	11	2.0	125	2	Q962W4	Q962w4 trypanosoma	222	11	2.0	277	2	Q86IC7	Q86ic7 dictyosteli
150	11	2.0	126	2	Q61021	Q61021 trypanosoma	223	11	2.0	278	2	Q25862	Q25862 plasmodium
151	11	2.0	126	2	Q61056	Q61056 trypanosoma	224	11	2.0	280	2	Q17639	Q17639 caenorhabdi
152	11	2.0	128	2	Q90603	Q90603 trypanosoma	225	11	2.0	283	2	Q86IL5	Q86il5 dictyosteli
153	11	2.0	128	2	Q9N1Q2	Q9niq2 trypanosoma	226	11	2.0	284	2	Q20202	Q20202 caenorhabdi
154	11	2.0	128	2	Q49334	Q49334 mycoplasma	227	11	2.0	293	2	Q25785	Q25785 plasmodium
155	11	2.0	131	2	Q9D9N0	Q9d9n0 mus musculu	228	11	2.0	295	2	Q9QYL6	Q9qyl6 mus musculu
156	11	2.0	136	2	Q86KRO	Q86kro dictyosteli	229	11	2.0	295	2	Q92Z88	Q92z88 mus musculu
157	11	2.0	139	2	Q61037	Q61037 trypanosoma	230	11	2.0	302	2	Q95U23	Q95u23 drosophila
158	11	2.0	139	2	Q90601	Q90601 trypanosoma	231	11	2.0	303	2	Q91LM9	Q91lm9 white spot
159	11	2.0	139	2	Q6WAZ8	Q6waz8 trypanosoma	232	11	2.0	304	1	YOOC_CABEL	Q99300 caenorhabdi
160	11	2.0	139	2	Q7Z0N2	Q7z0n2 caenorhabdi	233	11	2.0	306	2	Q9QV14	Q9qvl4 mus musculu
161	11	2.0	140	2	Q962W5	Q962w5 trypanosoma	234	11	2.0	312	2	Q01824	Q01824 pneumocysti
162	11	2.0	143	2	Q15776	Q15776 trypanosoma	235	11	2.0	312	2	Q6VZS5	Q6vzs5 canarypox v
163	11	2.0	143	2	Q81T82	Q81t82 plasmodium	236	11	2.0	320	2	Q9U448	Q9u448 dictyosteli
164	11	2.0	146	2	Q6RUB8	Q6rub8 chimpanzee	237	11	2.0	322	2	Q8SXG9	Q8sxg9 drosophila
165	11	2.0	147	2	Q6RUE3	Q6rue3 chimpanzee	238	11	2.0	323	2	Q86AP2	Q86ap2 dictyosteli
166	11	2.0	147	2	Q6RUE8	Q6rue8 chimpanzee	239	11	2.0	327	2	Q25334	Q25334 leishmania
167	11	2.0	148	2	Q61019	Q61019 trypanosoma	240	11	2.0	335	2	Q96U23	Q96u23 neurospora
168	11	2.0	148	2	Q6WB00	Q6wb00 trypanosoma	241	11	2.0	335	2	Q86HM2	Q86hm2 dictyosteli
169	11	2.0	148	2	Q6RUC1	Q6ruc1 chimpanzee	242	11	2.0	341	2	Q8IMS9	Q8ims9 drosophila
170	11	2.0	150	2	Q9BJQ6	Q9bjq6 plasmodium	243	11	2.0	344	2	Q76741	Q76741 dictyosteli
171	11	2.0	150	2	Q9BJQ7	Q9bjq7 plasmodium	244	11	2.0	350	2	Q7Q1R0	Q7q1r0 anopheles g
172	11	2.0	150	2	Q95QX3	Q9gqx3 plasmodium	245	11	2.0	362	2	Q7YUQ6	Q7yuq6 trypanosoma
173	11	2.0	150	2	Q9VTF1	Q9vtf1 drosophila	246	11	2.0	364	2	Q7S2R4	Q7s2p4 neurospora
174	11	2.0	155	2	Q6U5F5	Q6u5f5 plasmodium	247	11	2.0	364	2	Q7YUQ5	Q7yuq5 trypanosoma
175	11	2.0	157	2	Q25713	Q25713 plasmodium	248	11	2.0	364	2	Q688R2	Q688r2 oryza sativ
176	11	2.0	160	2	Q94669	Q94669 plasmodium	249	11	2.0	365	2	Q7YUV8	Q7yuv8 trypanosoma
177	11	2.0	164	2	Q9BJQ5	Q9bjq5 plasmodium	250	11	2.0	365	2	Q869R5	Q869r5 dictyosteli

251	11	2.0	366	2	Q9YZ40	Q9YZ40 chimpanzee	324	11	2.0	571	2	Q8NJP9	Q8NJP9 candida alb
252	11	2.0	367	2	Q7YU08	Q7YU08 trypanosoma	325	11	2.0	573	2	Q8TOR9	Q8TOR9 drosophila
253	11	2.0	369	2	Q7YUQ1	Q7YUQ1 trypanosoma	326	11	2.0	574	2	Q9DSX5	Q9DSX5 m mus muscu
254	11	2.0	369	2	Q7YUQ2	Q7YUQ2 trypanosoma	327	11	2.0	579	2	Q7SDE0	Q7SDE0 neurospora
255	11	2.0	369	2	Q7YUQ3	Q7YUQ3 trypanosoma	328	11	2.0	600	2	Q17490	Q17490 anopheles g
256	11	2.0	369	2	Q7YUQ4	Q7YUQ4 trypanosoma	329	11	2.0	600	2	Q7PV58	Q7PV58 anopheles g
257	11	2.0	374	2	Q7QCS5	Q7QCS5 anopheles g	330	11	2.0	602	2	Q86JM6	Q86JM6 dictyosteli
258	11	2.0	385	2	Q7LZ8R	Q7LZ8R ictalurus p	331	11	2.0	627	2	Q6UN14	Q6UN14 leishmania
259	11	2.0	386	2	Q01759	Q01759 pneumocysti	332	11	2.0	636	2	Q8MP77	Q8MP77 dictyosteli
260	11	2.0	386	2	Q9GZ9H	Q9GZ9H caenorhabdi	333	11	2.0	648	2	Q95QX0	Q95QX0 caenorhabdi
261	11	2.0	390	2	Q8MP08	Q8MP08 bombyx mori	334	11	2.0	657	2	Q86J19	Q86J19 dictyosteli
262	11	2.0	392	2	Q8IIC1	Q8IIC1 plasmodium	335	11	2.0	660	2	Q94485	Q94485 dictyosteli
263	11	2.0	392	2	Q69Z58	Q69Z58 mus musculu	336	11	2.0	667	2	Q7VYV0	Q7VYV0 cryptospori
264	11	2.0	398	2	Q22902	Q22902 caenorhabdi	337	11	2.0	680	2	Q93374	Q93374 caenorhabdi
265	11	2.0	399	2	Q86TX1	Q86TX1 dictyosteli	338	11	2.0	681	2	Q9XUS9	Q9XUS9 caenorhabdi
266	11	2.0	404	2	Q86SB6	Q86SB6 drosophila	339	11	2.0	684	2	Q8T848	Q8T848 dictyosteli
267	11	2.0	417	2	Q01760	Q01760 pneumocysti	340	11	2.0	689	2	Q9Y114	Q9Y114 haemochus
268	11	2.0	422	2	Q91BH8	Q91BH8 spodoptera	341	11	2.0	693	2	Q07241	Q07241 pneumocysti
269	11	2.0	425	2	Q15755	Q15755 dictyosteli	342	11	2.0	697	1	CRAC_DICDI	CRAC_DICDI
270	11	2.0	425	2	Q86AP8	Q86AP8 dictyosteli	343	11	2.0	701	2	Q9U233	Q9U233 caenorhabdi
271	11	2.0	425	2	Q9GZFS	Q9GZFS caenorhabdi	344	11	2.0	714	2	Q9NGW9	Q9NGW9 dictyosteli
272	11	2.0	426	2	Q961P7	Q961P7 drosophila	345	11	2.0	730	2	Q7QCW4	Q7QCW4 anopheles g
273	11	2.0	426	2	Q9EN59	Q9EN59 spodoptera	346	11	2.0	732	2	Q95Q40	Q95Q40 caenorhabdi
274	11	2.0	427	2	Q86I17	Q86I17 dictyosteli	347	11	2.0	732	2	Q86L25	Q86L25 dictyosteli
275	11	2.0	433	2	Q9VFD1	Q9VFD1 drosophila	348	11	2.0	734	2	Q86AK8	Q86AK8 dictyosteli
276	11	2.0	438	2	Q86UC6	Q86UC6 dictyosteli	349	11	2.0	737	2	Q86IX1	Q86IX1 dictyosteli
277	11	2.0	440	2	Q20659	Q20659 caenorhabdi	350	11	2.0	740	2	Q8T2A4	Q8T2A4 dictyosteli
278	11	2.0	442	2	Q9BY67	Q9BY67 homo sapien	351	11	2.0	742	2	Q23766	Q23766 chironomus
279	11	2.0	445	2	Q7Q956	Q7Q956 anopheles g	352	11	2.0	746	2	Q9V515	Q9V515 drosophila
280	11	2.0	448	2	Q8R4L1	Q8R4L1 mus musculu	353	11	2.0	748	2	Q44416	Q44416 chironomus
281	11	2.0	448	2	Q9N310	Q9N310 caenorhabdi	354	11	2.0	754	2	Q6JXF2	Q6JXF2 schistosoma
282	11	2.0	453	2	Q870S8	Q870S8 neurospora	355	11	2.0	756	2	Q86IG6	Q86IG6 dictyosteli
283	11	2.0	456	2	Q8R5M8	Q8R5M8 mus musculu	356	11	2.0	758	2	Q8MPP3	Q8MPP3 drosophila
284	11	2.0	465	2	Q86KH4	Q86KH4 dictyosteli	357	11	2.0	761	2	Q9VD08	Q9VD08 drosophila
285	11	2.0	467	1	Q55279	Q55279 avian adeno	358	11	2.0	769	2	Q17921	Q17921 caenorhabdi
286	11	2.0	468	2	Q83H_DICDI	Q83H_DICDI	359	11	2.0	771	2	Q86H13	Q86H13 dictyosteli
287	11	2.0	474	2	Q7S290	Q7S290 neurospora	360	11	2.0	781	2	Q52753	Q52753 ruminococcu
288	11	2.0	476	2	Q6AYP5	Q6AYP5 ratu	361	11	2.0	782	2	Q75J44	Q75J44 dictyosteli
289	11	2.0	477	2	Q23993	Q23993 drosophila	362	11	2.0	793	1	REGA_DICDI	REGA_DICDI
290	11	2.0	482	2	Q7SE49	Q7SE49 neurospora	363	11	2.0	795	2	Q9Y1J5	Q9Y1J5 dictyosteli
291	11	2.0	484	1	OAR1_LOCM1	OAR1_LOCM1	364	11	2.0	813	2	P90538	P90538 dictyosteli
292	11	2.0	484	1	OAR2_LOCM1	OAR2_LOCM1	365	11	2.0	814	2	Q7S1S5	Q7S1S5 neurospora
293	11	2.0	493	2	Q7S3X3	Q7S3X3 neurospora	366	11	2.0	823	2	Q75JE3	Q75JE3 dictyosteli
294	11	2.0	503	2	Q13269	Q13269 caenorhabdi	367	11	2.0	831	2	Q15756	Q15756 dictyosteli
295	11	2.0	504	2	Q6CCD0	Q6CCD0 yarrowia li	368	11	2.0	837	2	Q7QFG1	Q7QFG1 anopheles g
296	11	2.0	511	2	Q869L8	Q869L8 dictyosteli	369	11	2.0	839	2	Q8TIP1	Q8TIP1 dictyosteli
297	11	2.0	512	1	WR33_ARATH	WR33_ARATH	370	11	2.0	844	2	Q86AP1	Q86AP1 dictyosteli
298	11	2.0	513	2	Q6V4A2	Q6V4A2 drosophila	371	11	2.0	867	2	Q90X49	Q90X49 brachydanio
299	11	2.0	513	2	Q86P22	Q86P22 drosophila	372	11	2.0	868	2	Q86IC5	Q86IC5 dictyosteli
300	11	2.0	513	2	Q9W3K2	Q9W3K2 drosophila	373	11	2.0	872	2	P90523	P90523 dictyosteli
301	11	2.0	517	1	LA1C_DIACA	P27486 dianthus ca	374	11	2.0	874	2	Q76535	Q76535 dictyosteli
302	11	2.0	518	2	Q43753	Q43753 dianthus ca	375	11	2.0	887	2	Q23913	Q23913 dictyosteli
303	11	2.0	519	2	Q09592	Q09592 caenorhabdi	376	11	2.0	889	2	Q23895	Q23895 dictyosteli
304	11	2.0	520	2	Q871E3	Q871E3 neurospora	377	11	2.0	893	2	Q8T7V3	Q8T7V3 dictyosteli
305	11	2.0	520	2	Q9VHU5	Q9VHU5 drosophila	378	11	2.0	895	2	Q7KWT7	Q7KWT7 dictyosteli
306	11	2.0	522	2	Q75JX8	Q75JX8 dictyosteli	379	11	2.0	895	2	Q86A69	Q86A69 dictyosteli
307	11	2.0	524	2	Q653F4	Q653F4 oryza sativ	380	11	2.0	910	2	Q75JH3	Q75JH3 dictyosteli
308	11	2.0	525	1	NU62_RAT	P17955 rattus norv	381	11	2.0	911	2	Q7SKW1	Q7SKW1 neurospora
309	11	2.0	526	1	NU62_MOUSE	Q63850 mus musculu	382	11	2.0	919	2	Q86J16	Q86J16 dictyosteli
310	11	2.0	526	2	Q8VAG9	Q8VAG9 white spot	383	11	2.0	930	2	Q8T1S8	Q8T1S8 dictyosteli
311	11	2.0	527	2	Q8QTA8	Q8QTA8 white spot	384	11	2.0	931	2	Q86K99	Q86K99 dictyosteli
312	11	2.0	530	2	Q61209	Q61209 caenorhabdi	385	11	2.0	932	2	Q15924	Q15924 dictyosteli
313	11	2.0	534	2	Q7KTI5	Q7KTI5 drosophila	386	11	2.0	937	2	Q86L47	Q86L47 dictyosteli
314	11	2.0	538	2	Q75JS0	Q75JS0 dictyosteli	387	11	2.0	947	2	Q86H44	Q86H44 dictyosteli
315	11	2.0	542	2	Q966U0	Q966U0 dictyosteli	388	11	2.0	956	2	Q00908	Q00908 cryptospori
316	11	2.0	543	2	Q84U26	Q84U26 mesembryant	389	11	2.0	961	2	Q8IC12	Q8IC12 plasmodium
317	11	2.0	547	2	Q7PDW3	Q7PDW3 anopheles g	390	11	2.0	967	2	Q86AS0	Q86AS0 dictyosteli
318	11	2.0	548	2	Q7S905	Q7S905 neurospora	391	11	2.0	975	2	Q15743	Q15743 dictyosteli
319	11	2.0	551	1	PPBJ_RAT	P51740 rattus norv	392	11	2.0	984	2	Q9Y1P7	Q9Y1P7 cryptospori
320	11	2.0	552	2	Q19659	Q19659 caenorhabdi	393	11	2.0	992	2	Q86H50	Q86H50 dictyosteli
321	11	2.0	559	1	PPBI_MOUSE	P24822 mus musculu	394	11	2.0	995	2	Q95WL0	Q95WL0 lymantria d
322	11	2.0	559	2	Q9VN36	Q9VN36 drosophila	395	11	2.0	1000	2	Q81TD4	Q81TD4 helicoverpa
323	11	2.0	562	2	Q9JKS8	Q9JKS8 rattus norv	396	11	2.0				

397	11	2.0	1002	2	074674	074674 pneumocysti	470	11	2.0	1598	2	Q95YM8	Q95ym8 apis mellif
398	11	2.0	1010	2	086J70	086j70 dictyosteli	471	11	2.0	1635	2	017412	017412 aedes aegypt
399	11	2.0	1010	2	09U6W2	09u6w2 heliothis v	472	11	2.0	1641	2	02GRZ3	02grz3 caenorhabdi
400	11	2.0	1011	2	09NHZ9	09nhz9 helicoverpa	473	11	2.0	1670	2	023901	023901 dictyosteli
401	11	2.0	1013	2	08WSZ2	08wsz2 helicoverpa	474	11	2.0	1671	2	0869S5	0869s5 dictyosteli
402	11	2.0	1014	2	0962B4	0962b4 helicoverpa	475	11	2.0	1672	2	0869N0	0869n0 dictyosteli
403	11	2.0	1014	2	06UY54	06uy54 helicoverpa	476	11	2.0	1679	2	07KT16	07kt16 drosophila
404	11	2.0	1014	2	06UY55	06uy55 helicoverpa	477	11	2.0	1682	2	015054	015054 homo sapien
405	11	2.0	1015	2	086AG0	086ag0 dictyosteli	478	11	2.0	1693	2	086H41	086h41 dictyosteli
406	11	2.0	1024	2	086AG0	086ag0 dictyosteli	479	11	2.0	1723	2	086H41	086h41 dictyosteli
407	11	2.0	1028	2	086G91	086g91 dictyosteli	480	11	2.0	1728	2	086H41	086h41 dictyosteli
408	11	2.0	1037	2	086H18	086h18 dictyosteli	481	11	2.0	1737	2	09TW28	09tw28 dictyosteli
409	11	2.0	1051	2	086I42	086i42 dictyosteli	482	11	2.0	1781	2	086KX8	086kx8 dictyosteli
410	11	2.0	1056	2	09W294	09w294 drosophila	483	11	2.0	1804	2	08WPD7	08wpd7 ciona intes
411	11	2.0	1065	2	001828	001828 pneumocysti	484	11	2.0	1832	2	086S03	086s03 dictyosteli
412	11	2.0	1069	2	096V12	096v12 pneumocysti	485	11	2.0	1849	2	086L41	086l41 dictyosteli
413	11	2.0	1069	2	06AHV5	06ahv5 pneumocysti	486	11	2.0	1853	2	07KT96	07kt96 drosophila
414	11	2.0	1069	2	09XYP1	09xyp1 dictyosteli	487	11	2.0	1858	1	P3K2_DICDI	P3K2_DICDI
415	11	2.0	1070	2	078734	078734 pneumocysti	488	11	2.0	1873	2	06S003	06s003 dictyosteli
416	11	2.0	1071	2	096V10	096v10 pneumocysti	489	11	2.0	1893	2	09NKC9	09nkc9 drosophila
417	11	2.0	1076	2	001830	001830 pneumocysti	490	11	2.0	1955	2	086K52	086k52 dictyosteli
418	11	2.0	1077	2	06AHS1	06ahs1 pneumocysti	491	11	2.0	2015	2	09U5Y1	09u5y1 dictyosteli
419	11	2.0	1079	2	096V11	096v11 pneumocysti	492	11	2.0	2048	2	086JW3	086jw3 dictyosteli
420	11	2.0	1081	2	06AHT3	06ah3 pneumocysti	493	11	2.0	2084	2	086I03	086i03 dictyosteli
421	11	2.0	1083	2	Q12075	Q12075 pneumocysti	494	11	2.0	2102	2	086I03	086i03 dictyosteli
422	11	2.0	1084	2	086LA6	086la6 dictyosteli	495	11	2.0	2107	2	086M87	086m87 dictyosteli
423	11	2.0	1088	2	09VPC7	09vpc7 drosophila	496	11	2.0	2183	2	07KMS4	07kms4 dictyosteli
424	11	2.0	1089	2	044007	044007 dictyosteli	497	11	2.0	2208	2	086H41	086h41 dictyosteli
425	11	2.0	1092	2	09UVY2	09uvy2 pneumocysti	498	11	2.0	2284	2	086I66	086i66 dictyosteli
426	11	2.0	1093	1	P14K_DICDI	P14K_DICDI	499	11	2.0	2296	2	Q9VNE0	Q9vne0 drosophila
427	11	2.0	1099	1	PLC1_CANAL	PLC1_CANAL	500	11	2.0	2344	2	086H67	086h67 dictyosteli
428	11	2.0	1123	2	086K40	086k40 dictyosteli	501	11	2.0	2528	2	086J36	086j36 dictyosteli
429	11	2.0	1131	2	07S0U4	07s0u4 neurospora	502	11	2.0	2536	2	0817C4	0817c4 dictyosteli
430	11	2.0	1136	2	09V5A6	09v5a6 drosophila	503	11	2.0	2621	2	086JW3	086jw3 dictyosteli
431	11	2.0	1138	2	08IS12	08is12 dictyosteli	504	11	2.0	2625	2	086M29	086m29 dictyosteli
432	11	2.0	1155	2	078552	078552 neurospora	505	11	2.0	2650	2	086AF1	086af1 dictyosteli
433	11	2.0	1158	2	08IU40	08iu40 dictyosteli	506	11	2.0	2678	2	09ND94	09nda9 dictyosteli
434	11	2.0	1163	2	0869M3	0869m3 dictyosteli	507	11	2.0	2722	2	086JN0	086jn0 dictyosteli
435	11	2.0	1163	2	086IJ3	086ij3 dictyosteli	508	11	2.0	2843	2	0963L5	0963l5 dictyosteli
436	11	2.0	1166	2	08IP52	08ip52 drosophila	509	11	2.0	2994	2	Q95ZG5	Q95zg5 dictyosteli
437	11	2.0	1192	2	0869E1	0869e1 dictyosteli	510	11	2.0	3177	2	Q9NAL3	Q9nal3 dictyosteli
438	11	2.0	1212	2	086HH0	086hh0 dictyosteli	511	11	2.0	3295	2	066GT3	066gt3 rattus norv
439	11	2.0	1222	2	086HL6	086hl6 dictyosteli	512	11	2.0	3446	2	086AC8	086ac8 dictyosteli
440	11	2.0	1222	2	086SA4	086sa4 dictyosteli	513	11	2.0	3535	2	081C29	081c29 plasmodium
441	11	2.0	1224	2	07Q0G7	07q0g7 anopheles g	514	11	2.0	3550	2	066GT4	066gt4 rattus norv
442	11	2.0	1227	2	06T872	06t872 dictyosteli	515	11	2.0	6118	2	081396	081396 plasmodium
443	11	2.0	1240	2	09DWH8	09dwh8 rat cytomeg	516	11	2.0	6118	2	081396	081396 plasmodium
444	11	2.0	1264	2	086H34	086h34 dictyosteli	517	10	1.8	46	2	08J4I7	08j4i7 chimpanzee
445	11	2.0	1265	2	059920	059920 pneumocysti	518	10	1.8	78	2	Q26881	Q26881 trypanosoma
446	11	2.0	1271	1	Y338_MYCGE	Y338_MYCGE	519	10	1.8	86	2	07PFH5	07pfh5 anopheles g
447	11	2.0	1271	2	086I23	086i23 dictyosteli	520	10	1.8	113	2	061052	061052 trypanosoma
448	11	2.0	1285	2	07SGA3	07sga3 neurospora	521	10	1.8	120	2	Q9NIQ7	Q9niq7 trypanosoma
449	11	2.0	1298	2	086G47	086g47 dictyosteli	522	10	1.8	125	2	096VJ3	096vj3 pneumocysti
450	11	2.0	1303	2	066S84	066s84 okopileura	523	10	1.8	127	2	08LBJ3	08lbj3 arabidopsis
451	11	2.0	1307	2	086J04	086j04 dictyosteli	524	10	1.8	127	2	09LF22	09lf22 arabidopsis
452	11	2.0	1310	2	086I04	086i04 dictyosteli	525	10	1.8	134	2	P89278	P89278 xestia C-ni
453	11	2.0	1364	2	08T245	08t245 dictyosteli	526	10	1.8	137	2	Q8QHP0	Q8qhp0 human cytom
454	11	2.0	1380	2	086J05	086j05 dictyosteli	527	10	1.8	137	2	Q8QVD8	Q8qvd8 human cytom
455	11	2.0	1398	2	0869T5	0869t5 dictyosteli	528	10	1.8	137	2	Q8QVD8	Q8qvd8 human cytom
456	11	2.0	1402	2	096668	096668 dictyosteli	529	10	1.8	138	2	Q8QS26	Q8qs26 porcine her
457	11	2.0	1404	2	086I80	086i80 dictyosteli	530	10	1.8	143	2	061028	061028 trypanosoma
458	11	2.0	1405	2	07KWK0	07kwk0 dictyosteli	531	10	1.8	144	2	061KH6	061kh6 drosophila
459	11	2.0	1450	2	086A05	086a05 dictyosteli	532	10	1.8	146	2	06RUD1	06rud1 chimpanzee
460	11	2.0	1483	2	07KMU3	07kmu3 dictyosteli	533	10	1.8	146	2	Q8RUB5	Q8rub5 chimpanzee
461	11	2.0	1485	2	07KWP2	07kwp2 dictyosteli	534	10	1.8	146	2	Q6RUE0	Q6rue0 chimpanzee
462	11	2.0	1489	2	Q6NP07	Q6np07 drosophila	535	10	1.8	146	2	Q6RUE5	Q6rue5 chimpanzee
463	11	2.0	1513	2	Q7SCN1	Q7scn1 neurospora	536	10	1.8	147	2	06RUC3	06ruc3 chimpanzee
464	11	2.0	1518	2	086AN8	086an8 dictyosteli	537	10	1.8	147	2	Q6RUD3	Q6rud3 chimpanzee
465	11	2.0	1523	2	Q86L99	Q86l99 dictyosteli	538	10	1.8	147	2	Q6RUD8	Q6rud8 chimpanzee
466	11	2.0	1524	2	Q9V854	Q9v854 drosophila	539	10	1.8	147	2	Q6RUE9	Q6rue9 chimpanzee
467	11	2.0	1529	2	Q8IS21	Q8is21 dictyosteli	540	10	1.8	147	2	Q6RUF2	Q6ruf2 chimpanzee
468	11	2.0	1529	2	Q9GQC2	Q9gqc2 dictyosteli	541	10	1.8	148	2	Q6RUB6	Q6rub6 chimpanzee
469	11	2.0	1556	2	Q8ST25	Q8st25 dictyosteli	542	10	1.8	148	2	Q6RUE1	Q6rue1 chimpanzee

543	1.8	152	2	Q7S996	Q7s996 neurospora	616	10	1.8	630	2	Q22225	Q22225 caenorhabdi
544	1.8	157	2	Q7PRF6	Q7prf6 anopheles g	617	10	1.8	634	2	Q7KX02	Q7KX02 caenorhabdi
545	1.8	165	2	Q26877	Q26877 trypanosoma	618	10	1.8	648	1	KAPC_DICDI	P34099 dictyosteli
546	1.8	188	2	Q9NAQ1	Q9naq1 caenorhabdi	619	10	1.8	650	2	Q9TXB8	Q9txb8 dictyosteli
547	1.8	194	2	Q6IK11	Q6ik11 drosophila	620	10	1.8	650	1	ARE2_SACBA	Q87612 saccharomyc
548	1.8	194	2	Q7PIY5	Q7piy5 anopheles g	621	10	1.8	650	2	Q7SE54	Q7se54 neurospora
549	1.8	195	2	Q17630	Q17630 caenorhabdi	622	10	1.8	659	2	Q6ZBH4	Q6zbh4 oryza sativ
550	1.8	195	2	Q9W4V3	Q9w4v3 drosophila	623	10	1.8	662	1	MUC1_XENLA	Q05049 xenopus lae
551	1.8	197	2	Q26878	Q26878 trypanosoma	624	10	1.8	665	2	Q9N2V0	Q9n2v0 caenorhabdi
552	1.8	201	2	Q6WAZ6	Q6waz6 trypanosoma	625	10	1.8	696	2	Q86HT7	Q86ht7 dictyosteli
553	1.8	203	2	Q7S4M2	Q7s4m2 neurospora	626	10	1.8	699	2	Q9MAH1	Q9mah1 arabidopsis
554	1.8	206	2	Q7NDW9	Q7ndw9 gloeobacter	627	10	1.8	711	2	Q7S623	Q7s623 neurospora
555	1.8	207	2	Q26879	Q26879 trypanosoma	628	10	1.8	732	2	Q76NV3	Q76nv3 dictyosteli
556	1.8	211	2	Q00026	Q00026 ajilomyces	629	10	1.8	734	2	Q869R9	Q869r9 dictyosteli
557	1.8	211	2	Q6WAZ3	Q6waz3 trypanosoma	630	10	1.8	741	2	Q9PY56	Q9py56 xestia c-ni
558	1.8	219	2	Q7S8U5	Q7s8u5 neurospora	631	10	1.8	743	2	Q7PQ19	Q7pq19 anopheles g
559	1.8	225	2	Q7PK76	Q7pk76 anopheles g	632	10	1.8	752	2	Q8MNE2	Q8mne2 dictyosteli
560	1.8	227	2	Q7S8L4	Q7s8l4 neurospora	633	10	1.8	756	2	Q44417	Q44417 chironomus
561	1.8	231	2	Q9BIQ7	Q9biq7 cooperia pu	634	10	1.8	756	2	Q44418	Q44418 chironomus
562	1.8	253	2	Q6J981	Q6j981 arabidopsis	635	10	1.8	760	2	Q86A14	Q86a14 dictyosteli
563	1.8	254	2	Q7SEA4	Q7sea4 neurospora	636	10	1.8	761	2	Q75J89	Q75j89 dictyosteli
564	1.8	254	2	Q9W012	Q9w012 drosophila	637	10	1.8	788	2	Q9W180	Q9w180 drosophila
565	1.8	255	2	Q23112	Q23112 arabidopsis	638	10	1.8	791	1	CUDA_DICDI	O00841 dictyosteli
566	1.8	260	2	Q73T85	Q73t85 mycobacteri	639	10	1.8	799	2	Q7YVU4	Q7yvu4 cryptospori
567	1.8	263	1	SGS3_DROYA	P13728 drosophila	640	10	1.8	805	2	Q8MLP2	Q8mlp2 drosophila
568	1.8	272	2	Q86EW5	Q86ew5 schistosoma	641	10	1.8	812	2	Q86AW3	Q86aw3 dictyosteli
569	1.8	274	2	Q86KZ5	Q86kz5 dictyosteli	642	10	1.8	816	2	Q28331	Q28331 archaeoglob
570	1.8	275	2	Q8MNJ5	Q8mnj5 dictyosteli	643	10	1.8	816	2	Q86HD8	Q86hd8 dictyosteli
571	1.8	294	2	Q7PK72	Q7pk72 anopheles g	644	10	1.8	817	2	Q7QAN0	Q7qan0 anopheles g
572	1.8	307	1	SGS3_DROME	P02840 drosophila	645	10	1.8	820	2	Q754J6	Q754j6 ashbya gos
573	1.8	311	2	Q7O5Z0	Q7o5z0 anopheles g	646	10	1.8	828	2	Q75JJO	Q75jjo dictyosteli
574	1.8	335	2	Q9ZTC1	Q9ztc1 arabidopsis	647	10	1.8	835	2	Q75JY0	Q75jy0 dictyosteli
575	1.8	344	2	Q9N311	Q9n311 caenorhabdi	648	10	1.8	838	2	Q8R3S1	Q8r3s1 mus musculu
576	1.8	348	2	Q86JW4	Q86jw4 dictyosteli	649	10	1.8	853	2	Q7RWY1	Q7rwy1 neurospora
577	1.8	356	2	Q818P7	Q818p7 anopheles g	650	10	1.8	857	2	Q6C1I6	Q6cli6 varrowia li
578	1.8	356	2	Q7PG96	Q7pg96 anopheles g	651	10	1.8	888	2	Q6KAT0	Q6kat0 mus musculu
579	1.8	368	2	Q7PK74	Q7pk74 anopheles g	652	10	1.8	904	2	Q86HL7	Q86hl7 dictyosteli
580	1.8	391	2	Q7FAX5	Q7fxj5 anopheles g	653	10	1.8	908	2	Q80YR4	Q80yr4 mus musculu
581	1.8	393	2	Q13495	Q13495 neocallimas	654	10	1.8	911	2	Q86H89	Q86h89 dictyosteli
582	1.8	393	2	Q9ZVZ4	Q9zvz4 arabidopsis	655	10	1.8	919	2	Q91BB5	Q91bb5 spodoptera
583	1.8	395	2	Q9SNB6	Q9snb6 arabidopsis	656	10	1.8	921	2	Q869X3	Q869x3 dictyosteli
584	1.8	403	2	Q80UL4	Q8oul4 mus musculu	657	10	1.8	923	2	Q17970	Q17970 caenorhabdi
585	1.8	411	2	Q9F8S2	Q9f8s2 oryza sativ	658	10	1.8	941	2	Q869Y0	Q869y0 dictyosteli
586	1.8	413	2	Q7PK75	Q7pk75 anopheles g	659	10	1.8	947	2	Q45739	Q45739 dictyosteli
587	1.8	417	2	Q6V4B7	Q6v4b7 drosophila	660	10	1.8	970	2	Q86AL6	Q86al6 dictyosteli
588	1.8	429	2	Q7ZUE5	Q7zue5 brachydanio	661	10	1.8	979	2	Q7KWS3	Q7kws3 dictyosteli
589	1.8	431	2	Q22454	Q22454 caenorhabdi	662	10	1.8	990	2	Q75JK7	Q75jk7 dictyosteli
590	1.8	434	2	Q22458	Q22458 caenorhabdi	663	10	1.8	999	2	Q86UE5	Q86ue5 homo sapien
591	1.8	435	2	Q45851	Q45851 caenorhabdi	664	10	1.8	1008	2	Q74672	Q74672 pneumocysti
592	1.8	450	2	Q7S732	Q7s732 neurospora	665	10	1.8	1012	2	Q81I31	Q81i31 caenorhabdi
593	1.8	451	1	MYBH_DICDI	P34127 dictyosteli	666	10	1.8	1014	2	Q17505	Q17505 caenorhabdi
594	1.8	451	2	Q965P9	Q965p9 caenorhabdi	667	10	1.8	1017	2	Q74670	Q74670 pneumocysti
595	1.8	465	2	Q9BIP2	Q9bip2 cooperia pu	668	10	1.8	1022	2	Q74671	Q74671 pneumocysti
596	1.8	468	2	Q7XMQ1	Q7xmq1 oryza sativ	669	10	1.8	1026	2	Q74669	Q74669 pneumocysti
597	1.8	474	2	Q7JP78	Q7jp78 caenorhabdi	670	10	1.8	1028	2	Q74668	Q74668 pneumocysti
598	1.8	476	2	Q86J73	Q86j73 dictyosteli	671	10	1.8	1030	2	Q01829	Q01829 pneumocysti
599	1.8	480	2	Q7SHC6	Q7shc6 neurospora	672	10	1.8	1050	2	Q9U6U8	Q9u6u8 dictyosteli
600	1.8	480	2	Q873J7	Q873j7 neurospora	673	10	1.8	1059	2	Q7S4I6	Q7s4i6 neurospora
601	1.8	489	2	Q8MTX9	Q8mtx9 drosophila	674	10	1.8	1064	2	Q8N3D8	Q8n3d8 homo sapien
602	1.8	489	2	Q7Y187	Q7y187 oryza sativ	675	10	1.8	1080	2	Q6QR24	Q6qr24 trypanosoma
603	1.8	490	2	Q6V6N2	Q6v6n2 drosophila	676	10	1.8	1083	2	Q813G1	Q813g1 plasmodium
604	1.8	490	2	Q6V6N4	Q6v6n4 drosophila	677	10	1.8	1089	2	Q86KP4	Q86kp4 dictyosteli
605	1.8	490	2	Q6V6N6	Q6v6n6 drosophila	678	10	1.8	1109	2	Q8XMS8	Q8xms8 clostridium
606	1.8	522	2	Q86JZ2	Q86jz2 dictyosteli	679	10	1.8	1164	2	Q7YU46	Q7yu46 drosophila
607	1.8	525	2	Q8R3F1	Q8r3f1 mus musculu	680	10	1.8	1208	2	Q9NGW7	Q9ngw7 dictyosteli
608	1.8	547	2	Q81943	Q8i943 drosophila	681	10	1.8	1219	2	Q95XG5	Q95xg5 caenorhabdi
609	1.8	549	2	Q81Q59	Q8iqs9 drosophila	682	10	1.8	1278	2	Q7S2S7	Q7s2s7 neurospora
610	1.8	560	2	Q7P847	Q7pe47 anopheles g	683	10	1.8	1307	2	Q9V4J6	Q9v4j6 drosophila
611	1.8	579	2	Q8MT56	Q8mt56 drosophila	684	10	1.8	1325	2	Q75JAI	Q75jai dictyosteli
612	1.8	579	2	Q9VB46	Q9vb46 drosophila	685	10	1.8	1348	2	Q86KE2	Q86ke2 dictyosteli
613	1.8	581	2	Q81925	Q8i925 hyphantria	686	10	1.8	1377	2	O42695	O42695 candida alb
614	1.8	582	2	Q24160	Q24160 drosophila	687	10	1.8	1390	2	O77033	O77033 dictyosteli
615	1.8	626	2	Q9NDD1	Q9ndd1 leishmania	688	10	1.8	1418	2	Q86HT4	Q86ht4 dictyosteli

689	10	1.8	1457	2	Q81S20	Q81a20 dictyosteli	762	9	1.6	241	2	Q9VDR0	Q9vdr0 drosophila
690	10	1.8	1478	2	Q75J11	Q75j11 dictyosteli	763	9	1.6	245	2	Q872C8	Q872c8 neurospora
691	10	1.8	1541	2	Q9V8M7	Q9v8m7 drosophila	764	9	1.6	248	2	Q9BIQ6	Q9biq6 cooperia pu
692	10	1.8	1543	2	Q9GV71	Q9gv71 dictyosteli	765	9	1.6	248	2	Q9BIQ8	Q9biq8 cooperia pu
693	10	1.8	1553	2	Q95TR0	Q95tr0 drosophila	766	9	1.6	251	2	Q6CHT1	Q6cht1 yarowia li
694	10	1.8	1570	1	P3K1_DICD1	P54673 dictyosteli	767	9	1.6	252	2	Q8BQ75	Q8bq75 mus musculus
695	10	1.8	1616	2	Q7KUB3	Q7kub3 drosophila	768	9	1.6	257	2	Q8BQ75	Q8bq75 mus musculus
696	10	1.8	1616	2	Q9V5J0	Q9vej0 drosophila	769	9	1.6	261	2	Q6IH30	Q6ih30 drosophila
697	10	1.8	1629	2	Q75J33	Q75j33 dictyosteli	770	9	1.6	273	2	Q65XK9	Q65xk9 oryza sativ
698	10	1.8	1716	2	Q917Q8	Q917q8 drosophila	771	9	1.6	278	2	Q6FIQ7	Q6fiq7 candida gla
699	10	1.8	1726	2	Q6XHB2	Q6xbh2 dictyosteli	772	9	1.6	278	2	Q7Q5F5	Q7q5f5 anopheles g
700	10	1.8	1740	2	Q9HCJ0	Q9hcj0 homo sapien	773	9	1.6	279	2	Q7Q5F5	Q7q5f5 anopheles g
701	10	1.8	1761	2	Q6ZQH8	Q6zqh8 mus musculus	774	9	1.6	287	2	Q7Q5F5	Q7q5f5 anopheles g
702	10	1.8	1845	2	Q75I33	Q75i33 neurospora	775	9	1.6	287	2	Q7Q5F5	Q7q5f5 anopheles g
703	10	1.8	1880	2	Q8MP27	Q8mp27 dictyosteli	776	9	1.6	290	2	Q9Y1H4	Q9y1h4 dictyosteli
704	10	1.8	1887	2	Q8SSY6	Q8ssy6 dictyosteli	777	9	1.6	294	2	Q8C6E5	Q8c6e5 mus musculus
705	10	1.8	1912	2	Q9VS99	Q9vsa99 drosophila	778	9	1.6	308	2	Q9CSZ3	Q9csz3 mus musculus
706	10	1.8	1969	2	Q7SEZ5	Q7se25 neurospora	779	9	1.6	309	2	Q7S0P9	Q7s0p9 neurospora
707	10	1.8	2062	2	Q9SFP6	Q9spn6 dictyosteli	780	9	1.6	315	2	Q81QQ1	Q81qq1 drosophila
708	10	1.8	2112	2	Q61DH3	Q61dh3 drosophila	781	9	1.6	316	2	Q6FXF0	Q6fxf0 candida gla
709	10	1.8	2140	2	Q7SG46	Q7sg46 neurospora	782	9	1.6	316	2	Q6ZJ42	Q6zj42 oryza sativ
710	10	1.8	2181	2	Q9VRA6	Q9vra6 drosophila	783	9	1.6	318	2	Q9N4M7	Q9n4m7 caenorhabdi
711	10	1.8	2507	2	Q81TH8	Q81th8 dictyosteli	784	9	1.6	321	2	Q17977	Q17977 caenorhabdi
712	10	1.8	2546	2	Q6XHA6	Q6xha6 dictyosteli	785	9	1.6	327	2	Q717U3	Q717u3 nicotiana t
713	10	1.8	2800	2	Q6XHB1	Q6xnb1 dictyosteli	786	9	1.6	327	2	Q9SXX1	Q9sxx1 nicotiana t
714	10	1.8	2833	2	Q9VP13	Q9vp13 drosophila	787	9	1.6	327	2	Q8BY98	Q8by98 m mus muscu
715	10	1.8	2929	2	Q86JG5	Q86jg5 dictyosteli	788	9	1.6	332	1	P111_MOUSE	Q9c2x5 mus musculus
716	10	1.8	3135	2	Q7KHU2	Q7kuh2 drosophila	789	9	1.6	354	2	Q86HY8	Q86hy8 dictyosteli
717	10	1.8	3166	2	Q9W3Z0	Q9w3z0 drosophila	790	9	1.6	359	2	Q90351	Q90351 cornuix co
718	10	1.8	3455	2	Q6R5A9	Q6r5a9 tenebrio mo	791	9	1.6	364	2	Q9YZ52	Q9yz52 chimpanzee
719	10	1.8	4498	2	Q9W2Z3	Q9w2z3 drosophila	792	9	1.6	368	2	Q9YZ45	Q9yz45 chimpanzee
720	10	1.8	8905	2	Q7KR73	Q7kr73 drosophila	793	9	1.6	389	2	Q9V6R9	Q9v6r9 drosophila
721	9	1.6	41	2	Q7R9M4	Q7r9m4 plasmodium	794	9	1.6	390	2	Q6CEI2	Q6cei2 yarowia li
722	9	1.6	48	2	Q87172	Q87172 chimpanzee	795	9	1.6	390	2	Q81927	Q81927 hyphantria
723	9	1.6	77	2	Q86NZ8	Q86nz8 drosophila	796	9	1.6	396	2	Q8H880	Q8h880 oryza sativ
724	9	1.6	93	2	Q9VUA6	Q9vua6 drosophila	797	9	1.6	408	2	Q6W4E5	Q6w4e5 drosophila
725	9	1.6	105	2	Q7S3N9	Q7s3n9 neurospora	798	9	1.6	408	2	Q6W4E7	Q6w4e7 drosophila
726	9	1.6	106	2	Q9NIQ6	Q9niq6 trypanosoma	799	9	1.6	408	2	Q6W4E9	Q6w4e9 drosophila
727	9	1.6	109	2	Q9VTF2	Q9vtf2 drosophila	800	9	1.6	408	2	Q6W4F2	Q6w4f2 drosophila
728	9	1.6	117	2	Q9BJP9	Q9bjp9 plasmodium	801	9	1.6	408	2	Q6W4G1	Q6w4g1 drosophila
729	9	1.6	120	2	Q04420	Q04420 chimpanzee	802	9	1.6	408	2	Q6W4G4	Q6w4g4 drosophila
730	9	1.6	122	2	Q81PE4	Q81pe4 drosophila	803	9	1.6	408	2	Q6W4G7	Q6w4g7 drosophila
731	9	1.6	124	2	Q7RUG3	Q7rug3 neurospora	804	9	1.6	408	2	Q6W4H2	Q6w4h2 drosophila
732	9	1.6	127	2	Q61035	Q61035 trypanosoma	805	9	1.6	409	2	Q8MKM3	Q8mkm3 drosophila
733	9	1.6	131	2	Q61J65	Q61j65 drosophila	806	9	1.6	409	2	Q86JU3	Q86ju3 dictyosteli
734	9	1.6	135	2	Q01623	Q01623 pneumocysti	807	9	1.6	410	2	Q6Y144	Q6y144 lactuca sal
735	9	1.6	142	2	Q26943	Q26943 trypanosoma	808	9	1.6	413	2	Q86I25	Q86i25 dictyosteli
736	9	1.6	145	2	Q6RUB0	Q6rub0 chimpanzee	809	9	1.6	413	2	Q8Y126	Q8y126 lactuca sat
737	9	1.6	147	2	Q6RUD2	Q6rud2 chimpanzee	810	9	1.6	422	2	Q86AN7	Q86an7 dictyosteli
738	9	1.6	147	2	Q6RUE4	Q6rue4 chimpanzee	811	9	1.6	426	2	Q6Y159	Q6y159 lactuca sat
739	9	1.6	158	2	Q68KL8	Q68kl8 nasonia vit	812	9	1.6	433	1	RTC1_DICD1	Q15746 dictyosteli
740	9	1.6	166	2	Q73V43	Q73v43 mycobacteri	813	9	1.6	435	2	Q9FPA2	Q9fpa2 oryza sativ
741	9	1.6	168	2	Q61IG4	Q61ig4 drosophila	814	9	1.6	438	2	Q8LH48	Q8lh48 oryza sativ
742	9	1.6	170	2	Q8TIF7	Q8tif7 dictyosteli	815	9	1.6	440	2	Q7S0A4	Q7s0a4 neurospora
743	9	1.6	172	2	Q96QT0	Q96qt0 homo sapien	816	9	1.6	441	2	Q8T1W3	Q8t1w3 dictyosteli
744	9	1.6	178	2	Q922K0	Q922k0 mus musculus	817	9	1.6	444	2	Q6MVL1	Q6mvl1 neurospora
745	9	1.6	184	2	Q6K966	Q6k966 oryza sativ	818	9	1.6	453	2	Q12945	Q12945 gallus gall
746	9	1.6	209	2	Q6K966	Q6k966 oryza sativ	819	9	1.6	461	2	Q9FUL9	Q9ful9 zea mays (m
747	9	1.6	202	2	Q9N2S0	Q9n2s0 trypanosoma	820	9	1.6	464	2	Q871Q8	Q871q8 neurospora
748	9	1.6	202	2	Q9AZX9	Q9azx9 bacterioph	821	9	1.6	475	2	Q8MSD6	Q8msd6 drosophila
749	9	1.6	202	2	Q9CGQ0	Q9cgq0 lactococcus	822	9	1.6	479	2	Q870W9	Q870w9 neurospora
750	9	1.6	207	2	Q61045	Q61045 trypanosoma	823	9	1.6	483	2	Q8T8X4	Q8t8x4 drosophila
751	9	1.6	208	2	Q9U0A0	Q9u0a0 plasmodium	824	9	1.6	483	2	Q9VK10	Q9vk10 drosophila
752	9	1.6	216	2	Q6A279	Q6a279 haemophilus	825	9	1.6	484	2	Q7P2E5	Q7p2e5 anopheles g
753	9	1.6	209	2	Q6X857	Q6x857 spodoptera	826	9	1.6	484	2	Q87G55	Q87g55 anopheles g
754	9	1.6	220	2	Q7S6T5	Q7s6ts neurospora	827	9	1.6	488	1	UZIP_DRONE	Uzip379 drosophila
755	9	1.6	224	2	Q8LGR8	Q8lgr8 lycopersico	828	9	1.6	490	2	Q61S08	Q61s08 oryza sativ
756	9	1.6	224	2	Q8VZ00	Q8vz00 arabidopsis	829	9	1.6	492	1	MEC2_RAT	Q00566 rattus norv
757	9	1.6	232	2	Q91Q83	Q91q83 arabidopsis	830	9	1.6	492	2	Q6PBT9	Q6pbt9 brachydanio
758	9	1.6	232	2	Q8N148	Q8ni48 homo sapien	831	9	1.6	494	2	Q9VWU0	Q9vwu0 drosophila
759	9	1.6	232	2	Q27423	Q27423 drosophila	832	9	1.6	496	2	Q7S6G4	Q7s6g4 neurospora
760	9	1.6	233	2	Q9U0C2	Q9u0c2 plasmodium	833	9	1.6	497	2	Q7SHV7	Q7shv7 neurospora
761	9	1.6	237	2	Q81RY4	Q81ry4 drosophila	834	9	1.6	499	2	Q95QK9	Q95qk9 caenorhabdi

835	9	1.6	505	2	Q7SHM1	Q7ehm1 neurospora	908	9	1.6	906	1	YPG1_DICSP	P10511 dictyosteli
836	9	1.6	509	2	Q94988	Q94888 drosophila	909	9	1.6	908	2	Q7S9B2	Q7S9b2 neurospora
837	9	1.6	509	2	Q7KM08	Q7km08 drosophila	910	9	1.6	914	2	Q86K47	Q86k47 dictyosteli
838	9	1.6	511	2	Q9VZH5	Q9vzh5 drosophila	911	9	1.6	919	2	Q8QS30	Q8q30 pongine her
839	9	1.6	514	2	Q7S7J2	Q7s7j2 neurospora	912	9	1.6	920	2	Q86KS1	Q86ks1 dictyosteli
840	9	1.6	519	2	Q7YTR7	Q7ytr7 caenorhabdi	913	9	1.6	929	2	Q8SSQ3	Q8sq3 dictyosteli
841	9	1.6	522	2	Q9XTK9	Q9xtk9 drosophila	914	9	1.6	941	2	Q9VXA2	Q9vxa2 drosophila
842	9	1.6	523	2	Q8OS23	Q8os23 chimpanzee	915	9	1.6	943	1	RNGB_DICDI	Q7m389 dictyosteli
843	9	1.6	525	2	Q7SBB9	Q7sbb9 neurospora	916	9	1.6	950	2	Q8MQN5	Q8mqn5 drosophila
844	9	1.6	526	2	Q6NVT9	Q6nvty9 homo sapien	917	9	1.6	968	2	Q8IR41	Q8ir41 drosophila
845	9	1.6	526	2	Q9UK58	Q9uk58 homo sapien	918	9	1.6	980	2	Q75JK3	Q75jk3 dictyosteli
846	9	1.6	527	2	Q9V9J6	Q9v9j6 drosophila	919	9	1.6	989	1	PTP3_DICDI	P54637 dictyosteli
847	9	1.6	527	2	Q9RIQ2	Q9riq2 rattus norv	920	9	1.6	1004	2	Q8MP26	Q8mp26 dictyosteli
848	9	1.6	532	2	Q8R5H9	Q8r5h9 mus musculu	921	9	1.6	1014	2	Q8SSY2	Q8ssy2 dictyosteli
849	9	1.6	534	2	Q7Q3D3	Q7q3d3 anopheles g	922	9	1.6	1024	2	Q75JA0	Q75ja0 dictyosteli
850	9	1.6	535	2	Q95QK8	Q95qk8 caenorhabdi	923	9	1.6	1043	2	Q8SSW7	Q8ssw7 dictyosteli
851	9	1.6	537	2	Q8IFC6	Q8ifp6 drosophila	924	9	1.6	1055	2	Q75J96	Q75j96 dictyosteli
852	9	1.6	540	2	Q7PS94	Q7ps94 anopheles g	925	9	1.6	1093	2	Q6NCH8	Q6nch8 rhodopeudo
853	9	1.6	545	2	Q8TIB3	Q8tib3 dictyosteli	926	9	1.6	1135	2	Q8SSU8	Q8ssu8 dictyosteli
854	9	1.6	546	2	Q8SXG6	Q8sxg6 drosophila	927	9	1.6	1144	2	Q86JE2	Q86je2 dictyosteli
855	9	1.6	556	2	Q9VFW1	Q9vfw1 drosophila	928	9	1.6	1147	2	Q95PH8	Q95ph8 dictyosteli
856	9	1.6	567	1	CHI3_CANAL	P40954 candida alb	929	9	1.6	1148	2	Q7QGP4	Q7qgp4 dictyosteli
857	9	1.6	578	2	Q8BPP4	Q8bpp4 mus musculu	930	9	1.6	1163	2	Q8GAR1	Q8gar1 dictyosteli
858	9	1.6	583	2	Q6K5Q2	Q6k5q2 oryza sativ	931	9	1.6	1175	2	Q9VRL7	Q9vrl7 drosophila
859	9	1.6	584	2	Q8G9Z3	Q8g9z3 dictyosteli	932	9	1.6	1176	2	Q8G9T7	Q8g9t7 dictyosteli
860	9	1.6	584	2	Q8K480	Q8k480 mus musculu	933	9	1.6	1205	2	Q86JG9	Q86jg9 dictyosteli
861	9	1.6	584	2	Q75JZ1	Q75jz1 dictyosteli	934	9	1.6	1222	2	Q7PPC0	Q7ppc0 anopheles g
862	9	1.6	607	2	Q9VXZ2	Q9vxz2 drosophila	935	9	1.6	1223	2	Q7KWX7	Q7kwx7 dictyosteli
863	9	1.6	610	2	Q6Y0X6	Q6y0x6 mus musculu	936	9	1.6	1241	2	Q75JCO	Q75jco dictyosteli
864	9	1.6	615	2	Q84SM4	Q84sm4 oryza sativ	937	9	1.6	1259	2	Q8G9X4	Q8g9x4 dictyosteli
865	9	1.6	617	1	ESR1_ICTPU	Q9yhz7 ictalurus p	938	9	1.6	1324	2	Q8MMQ2	Q8mmq2 dictyosteli
866	9	1.6	619	2	Q76P26	Q76p26 dictyosteli	939	9	1.6	1329	2	Q75JPO	Q75jpo dictyosteli
867	9	1.6	622	2	Q7S9H1	Q7s9h1 neurospora	940	9	1.6	1369	1	NFAS_CHICK	Q42414 gallus gall
868	9	1.6	622	2	Q9P389	Q9p389 neurospora	941	9	1.6	1407	2	Q9VB65	Q9vb65 drosophila
869	9	1.6	622	2	Q6Y141	Q6y141 lactuca sat	942	9	1.6	1408	1	SERR_DROME	P18168 drosophila
870	9	1.6	624	2	Q19780	Q19780 caenorhabdi	943	9	1.6	1458	2	Q61802	Q61802 caenorhabdi
871	9	1.6	638	1	UBQ2_MOUSE	Q9qzm0 mus musculu	944	9	1.6	1464	2	Q61802	Q61802 caenorhabdi
872	9	1.6	639	2	Q865X5	Q865x5 dictyosteli	945	9	1.6	1470	2	Q9VYK5	Q9vyk5 drosophila
873	9	1.6	653	2	Q7S3H6	Q7s3h6 neurospora	946	9	1.6	1495	2	Q8GAU5	Q8gau5 dictyosteli
874	9	1.6	655	2	Q7KKR0	Q7kkr0 drosophila	947	9	1.6	1518	2	Q21442	Q21442 caenorhabdi
875	9	1.6	658	2	Q7PNH9	Q7pnh9 anopheles g	948	9	1.6	1542	2	Q9VE07	Q9ve07 drosophila
876	9	1.6	662	2	Q8MQZ9	Q8mqz9 drosophila	949	9	1.6	1557	2	Q76NT8	Q76nt8 dictyosteli
877	9	1.6	662	2	Q9VGD0	Q9vgd0 drosophila	950	9	1.6	1752	2	Q9AE52	Q9ae52 ruminococcu
878	9	1.6	665	2	Q86HZ3	Q86hz3 dictyosteli	951	9	1.6	1789	2	Q8TI45	Q8ti45 dictyosteli
879	9	1.6	670	2	Q6PW12	Q6pw12 gallus gall	952	9	1.6	1795	2	Q76894	Q76894 drosophila
880	9	1.6	679	2	Q69HN9	Q69hn9 ciona intes	953	9	1.6	1806	2	Q8G9R4	Q8g9r4 dictyosteli
881	9	1.6	682	2	Q6DJD2	Q6did2 xenopus lae	954	9	1.6	1837	2	Q8IKF1	Q8ikf1 plasmodium
882	9	1.6	704	2	Q86IW5	Q86iw5 dictyosteli	955	9	1.6	1837	2	Q8G9Q5	Q8g9q5 dictyosteli
883	9	1.6	708	2	Q22806	Q22806 caenorhabdi	956	9	1.6	1847	2	Q7R5Z0	Q7r5z0 giardia lam
884	9	1.6	709	2	Q86IX4	Q86ix4 dictyosteli	957	9	1.6	1997	2	Q8ISL6	Q8isl6 plasmodium
885	9	1.6	714	2	Q8T210	Q8t210 dictyosteli	958	9	1.6	2087	2	Q8MXL2	Q8mxl2 leishmania
886	9	1.6	714	2	Q75JH7	Q75jh7 dictyosteli	959	9	1.6	2335	2	Q7VYR5	Q7vyr5 cryptospori
887	9	1.6	740	2	Q8FLY9	Q8fly9 corynebacte	960	9	1.6	2602	2	Q9VZ45	Q9vz45 drosophila
888	9	1.6	741	2	Q86HM4	Q86hm4 dictyosteli	961	9	1.6	2691	2	Q8ILS2	Q8il82 plasmodium
889	9	1.6	745	2	Q86CR5	Q86cr5 dictyosteli	962	9	1.6	2752	2	Q6LFI9	Q6lfi9 plasmodium
890	9	1.6	745	2	Q86KJ5	Q86kj5 dictyosteli	963	9	1.6	3633	2	Q8IHL0	Q8ihl0 dictyosteli
891	9	1.6	759	2	Q05143	Q05143 ruminococcu	964	9	1.6	3672	1	LML2_CAEEEL	Q21313 caenorhabdi
892	9	1.6	760	2	Q6BS16	Q6bs16 dictyosteli	965	9	1.6	3704	2	P91904	P91904 caenorhabdi
893	9	1.6	762	2	Q86AG3	Q86ag3 dictyosteli	966	9	1.6	4118	2	Q8GTF2	Q8gtf2 dictyosteli
894	9	1.6	767	2	Q86KE5	Q86ke5 dictyosteli	967	9	1.6	4138	2	Q8ILI3	Q8ili3 plasmodium
895	9	1.6	777	2	Q7K7G5	Q7k7g5 drosophila	968	9	1.6	7524	2	Q6PEZ0	Q6pezo mus musculu
896	9	1.6	780	1	A4_TETPL	Q73683 tetraodon f	969	9	1.6	9234	2	Q7KTP5	Q7ktp5 drosophila
897	9	1.6	786	2	Q86L03	Q86l03 dictyosteli	970	8	1.4	32	2	Q9S442	Q9s442 neisseria m
898	9	1.6	787	2	Q817P3	Q817p3 dictyosteli	971	8	1.4	47	2	Q8J4H6	Q8j4h6 chimpanzee
899	9	1.6	794	2	Q7KU09	Q7ku09 drosophila	972	8	1.4	47	2	Q87205	Q87205 chimpanzee
900	9	1.6	798	2	Q75J86	Q75j86 dictyosteli	973	8	1.4	47	2	Q87209	Q87209 chimpanzee
901	9	1.6	799	2	Q86A08	Q86a08 dictyosteli	974	8	1.4	47	2	Q87260	Q87260 chimpanzee
902	9	1.6	801	2	Q9W3K6	Q9w3k6 drosophila	975	8	1.4	47	2	Q87261	Q87261 chimpanzee
903	9	1.6	811	2	Q86I20	Q86i20 dictyosteli	976	8	1.4	48	2	Q7KJB8	Q7kjb8 plasmodium
904	9	1.6	814	2	Q80UR5	Q80ur5 mus musculu	977	8	1.4	48	2	Q9N6A3	Q9n6a3 plasmodium
905	9	1.6	867	2	Q8S2E7	Q8sze7 drosophila	978	8	1.4	48	2	Q8J4I4	Q8j4i4 chimpanzee
906	9	1.6	876	2	Q75JU2	Q75ju2 dictyosteli	979	8	1.4	48	2	Q87164	Q87164 chimpanzee
907	9	1.6	879	2	Q76874	Q76874 drosophila	980	8	1.4	48	2	Q87176	Q87176 chimpanzee

881 8 1.4 48 2 Q87193 chimpanzee
 982 8 1.4 48 2 Q87194 chimpanzee
 983 8 1.4 48 2 Q87201 chimpanzee
 984 8 1.4 48 2 Q87222 chimpanzee
 985 8 1.4 48 2 Q87255 chimpanzee
 986 8 1.4 48 2 Q87256 chimpanzee
 987 8 1.4 48 2 Q87257 chimpanzee
 988 8 1.4 48 2 Q87258 chimpanzee
 989 8 1.4 48 2 Q87265 chimpanzee
 990 8 1.4 48 2 Q87266 chimpanzee
 991 8 1.4 48 2 Q87269 chimpanzee
 992 8 1.4 48 2 Q87274 chimpanzee
 993 8 1.4 50 2 Q9N120 plasmidium
 994 8 1.4 51 2 Q87187 chimpanzee
 995 8 1.4 51 2 Q87191 chimpanzee
 996 8 1.4 51 2 Q87202 chimpanzee
 997 8 1.4 51 2 Q87203 chimpanzee
 998 8 1.4 51 2 Q87213 chimpanzee
 999 8 1.4 51 2 Q87263 chimpanzee
 1000 8 1.4 51 2 Q87270 chimpanzee

ALIGNMENTS

RESULT 1
 Q7DB77
 ID Q7DB77 PRELIMINARY; PRT; 558 AA.
 AC Q7DB77 Q7A9Q1;
 DT 05-JUL-2004 (TREMBlrel. 27, Created)
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
 DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
 DE Putative translocated intimin receptor protein (Translocated intimin
 receptor Tir).
 GN Name=tir; OrderedLocusNames=EC84561, Z5112;
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=83334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
 RX MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamoudis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RA "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
 RL Nature 409:529-533(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / RIMD 0509952 / EHEC;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Lida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
 RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
 RA "Complete genome sequence of enterohaemorrhagic Escherichia coli
 O157:H7 and genomic comparison with a laboratory strain K-12.";
 RL DNA Res. 8:11-22(2001).
 DR EMBL; AE005595; AAG58825.1; -;
 DR EMBL; AF002566; BAB37984.1; -;
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0007155; P:cell adhesion; IEA.
 DR InterPro; IPR003536; Tir receptor.
 DR Pfam; PF07489; Tir receptor C; 1.
 DR Pfam; PF03549; Tir receptor M; 1.
 DR Pfam; PF07490; Tir receptor N; 1.
 DR PRINTS; PR01370; TERNINIMINR.
 SQ SEQUENCE 558 AA; 58022 MW; 99C417222D4B4AA1 CRC64;

Query Match 100.0%; Score 558; DB 2; Length 558;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 558; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MPIGNLGNPNVNSIIPAPPLPSQTGACAGRGOLINSTPLGSRALFTPVNSMADSGD 60
 Db 1 MPIGNLGNPNVNSIIPAPPLPSQTGACAGRGOLINSTPLGSRALFTPVNSMADSGD 60
 Qy 61 NRASDVPLGNPMLAAASEITLNDGFEVLHDHGPDLTLNRQIGSSVFRVETQEDGKHIA 120
 Db 61 NRASDVPLGNPMLAAASEITLNDGFEVLHDHGPDLTLNRQIGSSVFRVETQEDGKHIA 120
 Qy 121 VQONGVETSVLSDQEVARLQSIDPEGKDFVFTGGRGGAGHAMVTVASDITEARQIL 180
 Db 121 VQONGVETSVLSDQEVARLQSIDPEGKDFVFTGGRGGAGHAMVTVASDITEARQIL 180
 Qy 181 ELLPEKGTGSGKAGESKVGELRESNSGAENTTETQTSLSRSPDKLWALGTAT 240
 Db 181 ELLPEKGTGSGKAGESKVGELRESNSGAENTTETQTSLSRSPDKLWALGTAT 240
 Qy 241 GLIGLAATGIQVALALTPEDSPPTTDPDAASATETATRDQLTKEAFQNPDKVNI 300
 Db 241 GLIGLAATGIQVALALTPEDSPPTTDPDAASATETATRDQLTKEAFQNPDKVNI 300
 Qy 301 LGNAIPSGVLKDDVVANIEEQAKAAGEAKQQAENNAQAOKKYDEQAQKEELKVS 360
 Db 301 LGNAIPSGVLKDDVVANIEEQAKAAGEAKQQAENNAQAOKKYDEQAQKEELKVS 360
 Qy 361 AGYGLSGALILGGGIGVAVTAALHRKNQPVETTTTTTTTTTTTTSARTVENKPNNTPA 420
 Db 361 AGYGLSGALILGGGIGVAVTAALHRKNQPVETTTTTTTTTTTTTSARTVENKPNNTPA 420
 Qy 421 NVDPGSEDWTESRRSSMASTSTFFDTSSIGTVQNPYADVKTSLHDSQVPTSNSTSVQ 480
 Db 421 NVDPGSEDWTESRRSSMASTSTFFDTSSIGTVQNPYADVKTSLHDSQVPTSNSTSVQ 480
 Qy 481 NMGNTDSVYVSTIQHPPRDITDNGARLLGNPSAGIQSTYARLALSGGLRHDMGGLTGS 540
 Db 481 NMGNTDSVYVSTIQHPPRDITDNGARLLGNPSAGIQSTYARLALSGGLRHDMGGLTGS 540
 Qy 541 SAVNTSNPPAGSHRFV 558
 Db 541 SAVNTSNPPAGSHRFV 558
 RESULT 2
 Q9R396 PRELIMINARY; PRT; 558 AA.
 ID Q9R396
 AC Q9R396;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
 DE Translocated intimin receptor Tir (L0027).
 GN Name=tir;
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=86/24;
 RX MEDLINE=99242825; PubMed=10225900;
 RA DeVinney R., Stein M., Reinscheid D., Abe A., Ruschkowski S.,
 RA Finlay B.B.;
 RA "Enterohaemorrhagic Escherichia coli O157:H7 produces Tir, which is
 RT translocated to the host cell membrane but is not tyrosine
 RT phosphorylated.";
 RL Infect. Immun. 67:2389-2398(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC43895, and EDL933;
 RX MEDLINE=98339885; PubMed=9673266;
 RA Perna N.T., Mayhew G.F., Posfai G., Elliott S., Donnenberg M.S.,

RA Kaper J.B., Blattner P.R.;
 RT "Molecular evolution of a pathogenicity island from enterohemorrhagic
 RT Escherichia coli O157:H7", (1998).
 RL Infect. Immun. 66:3810-3817 (1998).
 DR EMBL; AF125993; AAD29391.1; -.
 DR EMBL; AF071034; AAC31506.1; -.
 DR PIR; A98199; A98199.
 DR PIR; E86045; B86045.
 DR HSSP; Q9KWH9; 1F02.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0007155; P:cell adhesion; IEA.
 DR InterPro; IPR003536; Tir_receptor.
 DR Pfam; PF07489; Tir_receptor_C; 1.
 DR Pfam; PF03549; Tir_receptor_M; 1.
 DR Pfam; PF07490; Tir_receptor_N; 1.
 DR PRINTS; PR01370; TRNSINTIMNR.
 KW Receptor.
 SQ SEQUENCE 558 AA; 58022 MW; 99C417222D4B4AA1 CRC64;

Query Match 100.0%; Score 558; DB 2; Length 558;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 558; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPIGNLGNPNVNNNSIPPAPLPSTQDAGGGRGQLINSTGPGSLRALFTPVNSMADSGD 60
 DB 1 MPIGNLGNPNVNNNSIPPAPLPSTQDAGGGRGQLINSTGPGSLRALFTPVNSMADSGD 60

QY 61 NRASDVPGLPVNPMLAAASEITLNDGFEVLHDGPDLTNLRQIGSSVFRVETQEDGKHIA 120
 DB 61 NRASDVPGLPVNPMLAAASEITLNDGFEVLHDGPDLTNLRQIGSSVFRVETQEDGKHIA 120

QY 121 VQQRNGVTSVLSDOEYARLOSIDPEGKDKFVFTGGRGAGHAMVTVASDITEARQRI 180
 DB 121 VQQRNGVTSVLSDOEYARLOSIDPEGKDKFVFTGGRGAGHAMVTVASDITEARQRI 180

QY 181 ELLEPKGTGESKGAGSGVGLRESNAGNTTETQSTSSLSRSDPKMLALGTAT 240
 DB 181 ELLEPKGTGESKGAGSGVGLRESNAGNTTETQSTSSLSRSDPKMLALGTAT 240

QY 241 GLIGLAATGIVQALALTPEPDSPTTDPDAAASATETATRDLTKEAFQNPQKNIDE 300
 DB 241 GLIGLAATGIVQALALTPEPDSPTTDPDAAASATETATRDLTKEAFQNPQKNIDE 300

QY 301 LGNAIPSGVLKDDVVANIEEQAKAAGEAKQAIENNAQAQKYDEQQAQKEELKVSSG 360
 DB 301 LGNAIPSGVLKDDVVANIEEQAKAAGEAKQAIENNAQAQKYDEQQAQKEELKVSSG 360

QY 361 AGYGLSGALILGGGIGVAVTAALHRKNQPVETTTTTTTTTTSARTVENKPNNTPAQ 420
 DB 361 AGYGLSGALILGGGIGVAVTAALHRKNQPVETTTTTTTTTTSARTVENKPNNTPAQ 420

QY 421 NVDTPGSEDTMESRRSSMASTSTFPDTSIGTVQNPYADVKTSLHDSQVPTSNSNTSVQ 480
 DB 421 NVDTPGSEDTMESRRSSMASTSTFPDTSIGTVQNPYADVKTSLHDSQVPTSNSNTSVQ 480

QY 481 NMGNTDSVYVYTIQHPRTDTNGARLLGNPSAGIOSTYARLALSGGLRHDMGGLTGGSN 540
 DB 481 NMGNTDSVYVYTIQHPRTDTNGARLLGNPSAGIOSTYARLALSGGLRHDMGGLTGGSN 540

QY 541 SAVNTSNPPAPGSHRFV 558
 DB 541 SAVNTSNPPAPGSHRFV 558

RESULT 3

O85506 PRELIMINARY; PRT; 558 AA.
 AC O85506
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Translocated intimin receptor Tir.
 GN Name:tir;

RESULT 4

O7BHL5 PRELIMINARY; PRT; 547 AA.
 AC O7BHL5
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DE 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Translocated intimin receptor Tir.
 OS Citrobacter rodentium.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 RN NCBI_TaxID=562;
 RP SEQUENCE FROM N.A.
 RC STRAIN=95SP2;
 RX MEDLINE=99003184; PubMed=9784578;
 RA Paton A.W., Manning P.A., Woodrow M.C., Paton J.C.;
 RT "Translocated intimin receptors (Tir) of Shiga-toxicogenic Escherichia
 RT coli isolates belonging to serogroups O26, O111, and O157 react with
 RT sera from patients with hemolytic-uremic syndrome and exhibit marked
 RT sequence heterogeneity";
 RL Infect. Immun. 66:5580-5586 (1998).
 DR EMBL; AF070067; AAC69314.1; -.
 DR HSSP; Q9KWH9; 1F02.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0007155; P:cell adhesion; IEA.
 DR InterPro; IPR003536; Tir_receptor.
 DR Pfam; PF07489; Tir_receptor_C; 1.
 DR Pfam; PF03549; Tir_receptor_M; 1.
 DR Pfam; PF07490; Tir_receptor_N; 1.
 DR PRINTS; PR01370; TRNSINTIMNR.
 KW Receptor.
 SQ SEQUENCE 558 AA; 58175 MW; CA2CDDAC94527C2B CRC64;

Query Match 60.9%; Score 340; DB 2; Length 558;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 440; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MPIGNLGNPNVNNNSIPPAPLPSTQDAGGGRGQLINSTGPGSLRALFTPVNSMADSGD 60
 DB 1 MPIGNLGNPNVNNNSIPPAPLPSTQDAGGGRGQLINSTGPGSLRALFTPVNSMADSGD 60

QY 61 NRASDVPGLPVNPMLAAASEITLNDGFEVLHDGPDLTNLRQIGSSVFRVETQEDGKHIA 120
 DB 61 NRASDVPGLPVNPMLAAASEITLNDGFEVLHDGPDLTNLRQIGSSVFRVETQEDGKHIA 120

QY 121 VQQRNGVTSVLSDOEYARLOSIDPEGKDKFVFTGGRGAGHAMVTVASDITEARQRI 180
 DB 121 VQQRNGVTSVLSDOEYARLOSIDPEGKDKFVFTGGRGAGHAMVTVASDITEARQRI 180

QY 181 ELLEPKGTGESKGAGSGVGLRESNAGNTTETQSTSSLSRSDPKMLALGTAT 240
 DB 181 ELLEPKGTGESKGAGSGVGLRESNAGNTTETQSTSSLSRSDPKMLALGTAT 240

QY 241 GLIGLAATGIVQALALTPEPDSPTTDPDAAASATETATRDLTKEAFQNPQKNIDE 300
 DB 241 GLIGLAATGIVQALALTPEPDSPTTDPDAAASATETATRDLTKEAFQNPQKNIDE 300

QY 301 LGNAIPSGVLKDDVVANIEEQAKAAGEAKQAIENNAQAQKYDEQQAQKEELKVSSG 360
 DB 301 LGNAIPSGVLKDDVVANIEEQAKAAGEAKQAIENNAQAQKYDEQQAQKEELKVSSG 360

QY 361 AGYGLSGALILGGGIGVAVTAALHRKNQPVETTTTTTTTTTSARTVENKPNNTPAQ 420
 DB 361 AGYGLSGALILGGGIGVAVTAALHRKNQPVETTTTTTTTTTSARTVENKPNNTPAQ 420

QY 421 NVDTPGSEDTMESRRSSMAST 441
 DB 421 NVDTPGSEDTMESRRSSMAST 441

Qy 281 DQLTKEAFQNPQNQKVNIDELGNAIPSG 308

```
DR PRINTS; PR01370; TRNSINTMINR.
KW Receptor. 538 AA; 55602 MW; 447052A0E3214D6D CRC64;
SQ SEQUENCE 538 AA; 55602 MW; 447052A0E3214D6D CRC64;

Query Match 3.9%; Score 22; DB 2; Length 538;
Best Local Similarity 100.0%; Pred. No. 1.6e-12;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 KVFVTGGRGGAGHAMVTVASDI 172
DB 152 KVFVTGGRGGAGHAMVTVASDI 173

RESULT 8
Q47014 PRELIMINARY; PRT; 538 AA.
AC Q47014;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Translocated intimin receptor tir (Translocated intimin co-receptor)
DE (EspE protein).
GN Name=tir; Synonym=espE;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=REPEC 84/110/1, and B65/56;
RA Krejany E.O.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=95ZG1;
RX MEDLINE=99003184; PubMed=9784578;
RA Paton A.W., Manning P.A., Woodrow M.C., Paton J.C.;
RT "Translocated intimin receptors (tir) of Shiga-toxicogenic Escherichia coli isolates belonging to serogroups O26, O111, and O157 react with sera from patients with hemolytic-uremic syndrome and exhibit marked sequence heterogeneity.";
RL Infect. Immun. 66:5580-5586 (1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=B10;
RX MEDLINE=20187493; PubMed=10722617;
RA DOI=10.1128/JAI.68.4.2171-2182.2000;
RA Marches O., Nougayrede J.P., Boullier S., Mainil J., Charlier G., Raymond I., Pohl P., Boury M., De Rycke J., Milon A., Oswald E.;
RT "Role of tir and intimin in the virulence of rabbit enteropathogenic Escherichia coli serotype O103:H2.";
RL Infect. Immun. 68:2171-2182 (2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=413/89-1;
RX MEDLINE=98294040; PubMed=9632251;
RA Deibel C., Kraemer S., Chakraborty T., Ebel P.;
RT "EspE, a novel secreted protein of attaching and effacing bacteria, is directly translocated into infected host cells where it appears as a tyrosine-phosphorylated 90 kDa protein.";
RL Mol. Microbiol. 28:463-474 (1998).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=413/89-1;
RX MEDLINE=98294040; PubMed=9632251;
RA Deibel C., Kraemer S., Chakraborty T., Ebel P.;
RT "EspE, a novel secreted protein of attaching and effacing bacteria, is directly translocated into infected host cells where it appears as a tyrosine-phosphorylated 90 kDa protein.";
RL Mol. Microbiol. 28:463-474 (1998).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=REPEC-1;
RX MEDLINE=21153569; PubMed=11254564;
RA DOI=10.1128/JAI.69.4.2107-2115.2001;
RA Zhu C., Agin T.S., Elliott S.J., Johnson L.A., Thate T.E., Kaper J.B., Boedeker E.C.;
RT "Complete nucleotide sequence and analysis of the locus of enterocyte
```

```
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003536; Tir receptor.
DR Pfam; PF07489; Tir_receptor_C; 1.
DR Pfam; PF03549; Tir_receptor_M; 1.
DR Pfam; PF07490; Tir_receptor_N; 1.
DR PRINTS; PR01370; TRNSINTMINR.
KW Receptor. 538 AA; 55420 MW; 31D7A8E227B3D06C CRC64;
SQ SEQUENCE 538 AA; 55420 MW; 31D7A8E227B3D06C CRC64;

Query Match 3.9%; Score 22; DB 2; Length 538;
Best Local Similarity 100.0%; Pred. No. 1.6e-12;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 KVFVTGGRGGAGHAMVTVASDI 172
DB 152 KVFVTGGRGGAGHAMVTVASDI 173

RESULT 9
Q47016 PRELIMINARY; PRT; 538 AA.
AC Q47016;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-AUG-1999 (TrEMBLrel. 11, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Translocated intimin receptor (Tir).
GN Name=tir;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=REPEC 83/39;
RA O'Brien R.A.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=REPEC 83/39;
RA Krejany E.O.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=REPEC-1;
RA Agin T.S., Boedeker E.C.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=REPEC-1;
RX MEDLINE=98254123; PubMed=9593291;
RA Elliott S.J., Mainwright L.A., McDaniel T.K., Jarvis K.G., Deng Y.K., Lai L.C., McNamara B.P., Donnerberg M.S., Kaper J.B.;
RT "The complete sequence of the locus of enterocyte effacement (LEE) from enteropathogenic Escherichia coli E2348/69.";
RL Mol. Microbiol. 28:1-4 (1998).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=REPEC-1;
RX MEDLINE=97055784; PubMed=8900070; DOI=10.1016/0378-1097(96)00371-0;
RA Agin T.S., Cantey J.R., Boedeker E.C., Wolf M.K.;
RT "Characterization of the eaeA gene from rabbit enteropathogenic Escherichia coli strain RDEC-1 and comparison to other eaeA genes from bacteria that cause attaching-effacing lesions.";
RL FEMS Microbiol. Lett. 144:249-258 (1996).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=REPEC-1;
RX MEDLINE=21153569; PubMed=11254564;
RA DOI=10.1128/JAI.69.4.2107-2115.2001;
RA Zhu C., Agin T.S., Elliott S.J., Johnson L.A., Thate T.E., Kaper J.B., Boedeker E.C.;
RT "Complete nucleotide sequence and analysis of the locus of enterocyte
```



```

DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003536; Tir_receptor.
DR Pfam; PF07489; Tir_receptor_C; 1.
DR Pfam; PF03549; Tir_receptor_M; 1.
DR Pfam; PF07490; Tir_receptor_N; 1.
DR PRINTS; PRO1370; TRNSINTMINR.
SQ SEQUENCE 550 AA; 56509 MW; 19DD08A9BE9251CB CRC64;

Query Match      2.5%; Score 14; DB 2; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.00016;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 IPPAPPLPSQTDGA 29
Db 16 IPPAPPLPSQTDGA 29

RESULT 13
O68258 PRELIMINARY; PRT; 551 AA.
AC O68258;
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Translocated intimin receptor.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98187918; PubMed=9529069;
RA Voss E., Paton A.W., Manning P.A., Paton J.C.;
RT "Molecular analysis of Shiga toxinigenic Escherichia coli O111:H-
RT proteins which react with sera from patients with hemolytic-uremic
RT syndrome.";
RL Infect. Immun. 66:1467-1472(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99003184; PubMed=9784578;
RA Paton A.W., Manning P.A., Woodrow M.C., Paton J.C.;
RT "Translocated intimin receptors (Tir) of Shiga-toxinigenic Escherichia
RT coli isolates belonging to serogroups O26, O111, and O157 react with
RT sera from patients with hemolytic-uremic syndrome and exhibit marked
RT sequence heterogeneity.";
RL Infect. Immun. 66:5580-5586(1998).
DR EMBL; AF025311; AAC69249.1; -.
DR HSP; O9KWH9; 1F02.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003536; Tir_receptor.
DR Pfam; PF07489; Tir_receptor_C; 1.
DR Pfam; PF03549; Tir_receptor_M; 1.
DR Pfam; PF07490; Tir_receptor_N; 1.
DR PRINTS; PRO1370; TRNSINTMINR.
SQ SEQUENCE 551 AA; 56975 MW; 6EC95F76EF0F44CC CRC64;

Query Match      2.5%; Score 14; DB 2; Length 551;
Best Local Similarity 100.0%; Pred. No. 0.00016;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 IPPAPPLPSQTDGA 29
Db 16 IPPAPPLPSQTDGA 29

RESULT 14
O9KWH9 PRELIMINARY; PRT; 552 AA.
AC O9KWH9;
DT 01-OCT-2000 (TREMBlrel. 15, Created)

us-09-189-415b-11.oligo.rup

```

```

DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Tir.
GN Name=tir;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HK01;
RA Abe A., Nagano H.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB036053; BAA96815.1; -.
DR PDB; 1F02; X-ray; T=271-336.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003536; Tir_receptor.
DR Pfam; PF07489; Tir_receptor_C; 1.
DR Pfam; PF03549; Tir_receptor_M; 1.
DR Pfam; PF07490; Tir_receptor_N; 1.
DR PRINTS; PRO1370; TRNSINTMINR.
SQ SEQUENCE 552 AA; 57005 MW; A8D79BE22EE50A4B CRC64;

Query Match      2.5%; Score 14; DB 2; Length 552;
Best Local Similarity 100.0%; Pred. No. 0.00016;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 IPPAPPLPSQTDGA 29
Db 16 IPPAPPLPSQTDGA 29

RESULT 15
Q6GED0 PRELIMINARY; PRT; 163 AA.
AC Q6GED0;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Putative exported protein.
GN OrderedLococusNames=SA2388;
OS Staphylococcus aureus (strain MRSA252).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=282456;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15213324; DOI=10.1073/pnas.0402521101;
RA Holden M.T.G., Feil E.J., Lindey J.A., Peacock S.J., Day N.P.J.,
RA Enright M.C., Foster T.J., Moore C.E., Hurst L., Atkin R., Barron A.,
RA Bason N., Bentley S.D., Chillingworth C., Chillingworth T.,
RA Churcher C., Clark L., Corton C., Cronin A., Doggett J., Dowd L.,
RA Feltwell T., Hance Z., Harris B., Hauser H., Holroyd S., Jagels K.,
RA James K.D., Lennard N., Line A., Mayes R., Moule S., Mungall K.,
RA Ormond D., Quail M.A., Rabinowitch E., Rutherford K.M., Sanders M.,
RA Sharp S., Simmonds M., Stevens K., Whitehead S., Barrell B.G.,
RA Spratt B.G., Parkhill J.;
RT "Complete genomes of two clinical Staphylococcus aureus strains:
RT evidence for the rapid evolution of virulence and drug resistance.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).
DR EMBL; EX571856; CAG41369.1; -.
DR InterPro; IPR007921; CHAP.
DR PROSITE; PS50911; CHAP; 1.
KW Complete proteome.
SQ SEQUENCE 163 AA; 17007 MW; 202B20997858FCC1 CRC64;

Query Match      2.3%; Score 13; DB 2; Length 163;
Best Local Similarity 100.0%; Pred. No. 0.00054;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 392 QTTTTTTTTTTTS 404
Db 40 QTTTTTTTTTTTS 52

```

Search completed: May 13, 2005, 12:06:49
Job time : 203 secs

```

US-09-346-374-243
; Sequence 243, Application US/09946374
; Publication No. US20030073129A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Tra
; TITLE OF INVENTION: Acids Encoding
; FILE REFERENCE: P2830P1C1
; CURRENT APPLICATION NUMBER: US/09/94
; CURRENT FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: 60/098716
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098723
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098749

```

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

; PRIOR FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: 60/104257
; PRIOR FILING DATE: 1998-10-14
; PRIOR APPLICATION NUMBER: 60/104987
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: 60/105000
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: 60/105002
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: 60/105104
; PRIOR FILING DATE: 1998-10-21
; PRIOR APPLICATION NUMBER: 60/105169
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: 60/105266
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: 60/105693
; PRIOR FILING DATE: 1998-10-26
; PRIOR APPLICATION NUMBER: 60/105694
; PRIOR FILING DATE: 1998-10-26
; PRIOR APPLICATION NUMBER: 60/105807

Query Match 6.3%; Score 180; DB 10; Length 596;
Best Local Similarity 19.2%; Pred. No. 0.0021;
Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;

QY 4 GNLGHPNVNNSIPAPPPLPSQTDGAGGRGQLINSTGPGSRALFTPVNSWADSGDNRA 63
Db 13 GLLLHLEAATNS-----NETSANTGSSVSSG-----ASTATNSG 49
QY 64 SDVPGPLVNPMLRAASEITLNDQFEVLHDHGLDNLNRIGSSV-PRVETQEDGKHIAGV 122
Db 50 SSVTSSGVSTATISGSSVTSN-GVSIV-TNSEPHHTSSGISTATNSEFTASGSIATN 107
QY 123 QRNGVETS--VVLSDQFYARLOSIDPEGKDKFVFTGCGRGAGHAMVTVASDITEARQIL 180
Db 108 SESSTTSSGASTATNSE-----SSTPSSGASTVNSGSSVTSNGASTATNSSESSTVSSRAS 163
QY 181 ELLEPKGTGESKAGESKGVGELRESNGAENTTETQSTSTSSLSRSDPKLWMLALCTVAT 240
Db 164 TATNSESSTLSSGASTATN-SDSSTTSSGASTATNSESSTSSG-----ASTAT 211
QY 241 GLIGLAATGIVQALALTPEDSPSTTTPDAAASATETATRDOLTKEAFQNPDKNVIDE 300
Db 212 N-----SESTVSSRASATNSESSTT---SSGASTATNSESRTNGAGTATNSESSTTS 264
QY 301 LG-----NAIPSGVLKDDVYANIEBQAKAAGEAKQQAIEENNAQAQKYEQQAKR 351
Db 265 SGASTATNSDSSVSSGA---STATNSESSTTSSGAST-----ATN 302
QY 352 QBELKVSSGAGYGL-SCALILGGIGIVATVAAHLRKNQVPVEQTTTTTTTTTSARTVEN 410
Db 303 SESSTTSSGASTATNSDSSSTSSGASTATNSESSTVSSGISTVTSSESSTPSSGANTATN 362
QY 411 KPANTTPACQNVDTPGSEDTMESRRSSMASTSTFTDTSIGTVQNPYADVKT-----464
Db 363 SESSTTSSGANTAT-NSESSTVSSGASTATNSESSTTSSGVTATNSESSTTSSGASTAT 421
QY 465 LHDQVPTNSNTSVQNMGTDSVVYSTIQHPRDTTNGARLLGNPSAGIQSTVARLAL 524
Db 422 NSDSSSTTSEASTATNSESSTVSSGISTVTSSESSTTSSGANTATNSESSTVSS-----474
QY 525 SGLRHDMLGGLTGGNSAVNTNPPAPG 553
Db 475 AGSGTAALATGMMHTSHSA-STAVSEAKPG 502

RESULT 2
US-10-006-867-100
; Sequence 100, Application US/10006867
; Publication No. US20020119130A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen

; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230RIC1
; CURRENT APPLICATION NUMBER: US/10/006,867
; CURRENT FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: 60/063435
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/064215
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088740
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088811
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088825
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088863
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090688
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091628
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/096012
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: 60/096757
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 60/096949
; PRIOR FILING DATE: 1998-08-18
; PRIOR APPLICATION NUMBER: 60/096959
; PRIOR FILING DATE: 1998-08-18
; PRIOR APPLICATION NUMBER: 60/097954
; PRIOR FILING DATE: 1998-08-26
; PRIOR APPLICATION NUMBER: 60/097971

;
; PRIOR FILING DATE: 1998-08-26
; PRIOR APPLICATION NUMBER: 60/097979
; PRIOR FILING DATE: 1998-08-26
; PRIOR APPLICATION NUMBER: 60/098749
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/099741
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099763
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099792
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099812
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099815
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/100627
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: 60/100662
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: 60/100683
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/100684
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/100930
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/101279
; PRIOR FILING DATE: 1998-09-22
; PRIOR APPLICATION NUMBER: 60/101475
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 60/101738
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101743
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101916
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/102570
; PRIOR FILING DATE: 1998-09-30
; PRIOR APPLICATION NUMBER: 60/103449
; PRIOR FILING DATE: 1998-10-06
; PRIOR APPLICATION NUMBER: 60/103678
; PRIOR FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: 60/103679
; PRIOR FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: 60/103711
; PRIOR FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: 60/105000
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: 60/105002
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: 60/105981
; PRIOR FILING DATE: 1998-10-27
; PRIOR APPLICATION NUMBER: 60/106030
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: 60/106464
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 60/106956
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: 60/108807
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: 60/112419
; PRIOR FILING DATE: 1998-12-15
; PRIOR APPLICATION NUMBER: 60/112422
; PRIOR FILING DATE: 1998-12-15
; PRIOR APPLICATION NUMBER: 60/112853
; PRIOR FILING DATE: 1998-12-16
; PRIOR APPLICATION NUMBER: 60/113011
; PRIOR FILING DATE: 1998-12-16
; PRIOR APPLICATION NUMBER: 60/112854
; PRIOR FILING DATE: 1998-12-16
; PRIOR APPLICATION NUMBER: 60/113300
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: 60/113408
; PRIOR FILING DATE: 1998-12-22

;
; PRIOR APPLICATION NUMBER: 60/113430
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/113621
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/114223
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 60/115614
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/116527
; PRIOR FILING DATE: 1999-01-20
; PRIOR APPLICATION NUMBER: 60/116843
; PRIOR FILING DATE: 1999-01-22
; PRIOR APPLICATION NUMBER: 60/119285
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: 60/119287
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: 60/119525
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: 60/119549
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: 60/120014
; PRIOR FILING DATE: 1999-02-11
; PRIOR APPLICATION NUMBER: 60/129122
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 60/129674
; PRIOR FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: 60/131291
; PRIOR FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: 60/138387
; PRIOR FILING DATE: 1999-06-09
; PRIOR APPLICATION NUMBER: 60/144791
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/169495
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: 60/175481
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 60/191007
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/199397
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/380139
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: 09/311832
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 09/380137
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: 09/380138
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: 09/380142

Query Match 6.3%; Score 180; DB 13; Length 596;

Best Local Similarity 19.2%; Pred. No. 0.0021;

Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;

Qy 4 GNLGHNPNVNSIPPAPPLPQTDGAGRGQLINSTGLSGRALFTPVNRNSWADSGDNRA 63

Db 13 GLLHLEAATNS-----NETSTANTSGSVISSG-----ASTATNSG 49

Qy 64 SDVPGLPVNPNRLAASEITLNDGFEVLHDHGPDLTLNRQIGSSV-FRVETQEDGKHIAG 122

Db 50 SSVTSSGVSTATISGSSVTSN-GVSIV-TNSEFHTTSSGISTATNSESTASSGSIATN 107

Qy 123 QRNGVETS--VVLSDQEVARLQSIDPEGKDFVFTGGRGGAGHAMVTVASDITEARQRIL 180

Db 108 SESSTTSSGASTATNSE-----SSTPSSGASTVTHSGSVTSSGASTATNSESTVSSRAS 163

Qy 181 ELLEPKGTGSKGAGESKGVGELRESNAGBNTTETQTSTSSLRSDPKLWLALGTVAT 240

Db 164 TATNSESSLSSGASTATN-SDSSTTSSGASTATNSESTTSSG-----ASTAT 211

Qy 241 GLIGLATGIQVALTPEPDSPTTDPDAAASATEATRDQLIKEAFQNPDKNQVNDI 300

Db 212 N-----SESSTVSSRASTATNSESTT---SSGASTATNSESTTSSNGAGTATNSESTTS 264

QY 301 LG-----NAIPSGVLKDDVVANIEQAKAAGEAKQQAIIENNAQAKKYDEQOAKR 351
Db 265 SGASTATNSDSTVSSGA---STATNSESTTSSGAST-----ATN 302
QY 352 QBELKVSSGAGYGL-SGALILGGIGGVAVTAALHRKNQPVQETTTTTTTTTTTSARTVEN 410
Db 303 SSSSTSSGASTATNSDSTSSGAGTATNSESTVSSGISTVTNSESTPSSGANTATN 362
QY 411 KPANNTPAGNVDTFGSEPTMESRRSSMASTSTFTDTSIGTVQNPYADVKT-----464
Db 363 SSSSTSSGANTAT-NSSESTVSSGASTATNSESTTSSGVTSTATNSESTTSSGASTAT 421
QY 465 LHDSQVPSNTSVNQMGNTDSVYVSTIQHPRTTDTNGALLGNPSAGIOSTVARIAL 524
Db 422 NSDSTSSSEASTATNSESTVSSGISTVTNSESTTSSGANTATNSGSSVTS-----474
QY 525 SGGLRHDGMLTGGNSAVNTSNNPPAPG 553
Db 475 AGSGTAALTGMTTSHSA-STAVSEAKPG 502

RESULT 3
US-10-052-586-310
; Sequence 310, Application US/10052586
; Publication No. US20020127584A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C1
; CURRENT APPLICATION NUMBER: US/10/052,586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063564
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063734
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063870
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066120
; PRIOR FILING DATE: 1997-11-21

; PRIOR APPLICATION NUMBER: 60/066466
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066772
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/069335
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069425
; PRIOR FILING DATE: 1997-12-12
; PRIOR APPLICATION NUMBER: 60/069870
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/068017
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579


```

; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-063-551-100

Query Match      6.3%; Score 180; DB 13; Length 596;
Best Local Similarity 19.2%; Pred. No. 0.0021;
Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;

QY 4 GNLGHNPNVNNISIPAPPPLPSQTDGAGRGQLINSTGLSGRALFTPPVNSMADSGDNRA 63
Db 13 GLLHLEAATNS-----NETSTANTGSSVSSG-----ASTATNSG 49
QY 64 SDVPGLPVNPMLAASEITLNDGFEVLHDHGPLDLNROIGSSV-PRVETQEDGKHIAVG 122
Db 50 SVTSSGVSTATISGSSVTSN-GVSIV-TNSEFHTTSSGISTATNSEFSTASSGISAIYN 107
QY 123 QNNGVETS--VVLSDOEYARLQSIDPEGKDKFVFTGGRGAGHAMVTVASDITEARQIL 180
Db 108 SSSSTSSGASTATNSE-----SSTPSSGASTVTVNSGSSVTSSGASTATNSESTVSSRAS 163
QY 181 ELLEPKGTGESKGAGSKGVGELRESNGAENTTQTSTSTSSLSRSDPKLMLALGTVAT 240
Db 164 TATNSESSTLSSGASTATN-SDSSTTSSGASTATNSESTTSSG-----ASTAT 211
QY 241 GLIGLAATGIVOALALTPEPDSPTTDDPAAASATETATRDQLTKEAFQNPQNKVNIDE 300
Db 212 N-----SESSTVSSRSTATNSESTT---SSGASTATNSESTTNGAGTATNSESTTS 264
QY 301 LG-----NAIPSGVLKDDVANIIEQAKAAGEAKQQAIIENNAQAQKYDEQOAKR 351
Db 265 SGASTATNSDSSTVSSGA---STATNSESTTSSGAST-----ATN 302
QY 352 QBELKVSGAGYGL-SGALILGGGIGVAVTAALHRKNQPVQOTTTTTTTTTTSTARTVEN 410
Db 303 SESSTTSSGASTATNSDSSTTSSGAGTATNSESTVSSGISTVTVNSSESSTPSSGANTATN 362
QY 411 KPANNTPAQGNVDTPGSEDPTMESRRSSMASTSTFTDTSIGTVQNPYADVTS-----464
Db 363 SESSTTSSGANTAT-NSESSTVSSGASTATNSESTTSSGVTATNSESTTSSGASTAT 421
QY 465 LHDSQVPTNSNTSVQNMGNDSVVYSTIQHPRTDTTNGARLLGNPSAGIOSTVARLAL 524
Db 422 NDSSTTSSEASTATNSESTVSSGISTVTVNSSESSTTSSGANTATNSGSSVTS-----474
QY 525 SGGLRDMGGLTGGNSAVNTSNPPAPG 553
Db 475 AGSGTAALTGMHTTSHSA-STAVSEAKPG 502

RESULT 5
US-10-063-551-100
; Sequence 100, Application US/10063551
; Publication No. US20020183494A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,551
; PRIOR FILING DATE: 2002-05-02
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 100
; LENGTH: 596

; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-063-551-100

Query Match      6.3%; Score 180; DB 13; Length 596;
Best Local Similarity 19.2%; Pred. No. 0.0021;
Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;

QY 4 GNLGHNPNVNNISIPAPPPLPSQTDGAGRGQLINSTGLSGRALFTPPVNSMADSGDNRA 63
Db 13 GLLHLEAATNS-----NETSTANTGSSVSSG-----ASTATNSG 49
QY 64 SDVPGLPVNPMLAASEITLNDGFEVLHDHGPLDLNROIGSSV-PRVETQEDGKHIAVG 122
Db 50 SVTSSGVSTATISGSSVTSN-GVSIV-TNSEFHTTSSGISTATNSEFSTASSGISAIYN 107
QY 123 QNNGVETS--VVLSDOEYARLQSIDPEGKDKFVFTGGRGAGHAMVTVASDITEARQIL 180
Db 108 SSSSTSSGASTATNSE-----SSTPSSGASTVTVNSGSSVTSSGASTATNSESTVSSRAS 163
QY 181 ELLEPKGTGESKGAGSKGVGELRESNGAENTTQTSTSTSSLSRSDPKLMLALGTVAT 240
Db 164 TATNSESSTLSSGASTATN-SDSSTTSSGASTATNSESTTSSG-----ASTAT 211
QY 241 GLIGLAATGIVOALALTPEPDSPTTDDPAAASATETATRDQLTKEAFQNPQNKVNIDE 300
Db 212 N-----SESSTVSSRSTATNSESTT---SSGASTATNSESTTNGAGTATNSESTTS 264
QY 301 LG-----NAIPSGVLKDDVANIIEQAKAAGEAKQQAIIENNAQAQKYDEQOAKR 351
Db 265 SGASTATNSDSSTVSSGA---STATNSESTTSSGAST-----ATN 302
QY 352 QBELKVSGAGYGL-SGALILGGGIGVAVTAALHRKNQPVQOTTTTTTTTTTSTARTVEN 410
Db 303 SESSTTSSGASTATNSDSSTTSSGAGTATNSESTVSSGISTVTVNSSESSTPSSGANTATN 362
QY 411 KPANNTPAQGNVDTPGSEDPTMESRRSSMASTSTFTDTSIGTVQNPYADVTS-----464
Db 363 SESSTTSSGANTAT-NSESSTVSSGASTATNSESTTSSGVTATNSESTTSSGASTAT 421
QY 465 LHDSQVPTNSNTSVQNMGNDSVVYSTIQHPRTDTTNGARLLGNPSAGIOSTVARLAL 524
Db 422 NDSSTTSSEASTATNSESTVSSGISTVTVNSSESSTTSSGANTATNSGSSVTS-----474
QY 525 SGGLRDMGGLTGGNSAVNTSNPPAPG 553
Db 475 AGSGTAALTGMHTTSHSA-STAVSEAKPG 502

RESULT 6
US-10-174-590-310
; Sequence 310, Application US/10174590
; Publication No. US20030008352A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C42
; CURRENT APPLICATION NUMBER: US/10/174,590
; CURRENT FILING DATE: 2002-06-18
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310
```

```
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-590-310

Query Match
Best Local Similarity 19.2%; Score 180; DB 14; Length 596;
Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;

QY 4 GNLGHNPNVNSIPAPPPLPSQTDGAGRGQLINSTGPLGSRALFTPVNRNSMADSGDNRA 63
Db 13 GLLHLEAATNS-----NETSANTGSSVSSG-----ASTATNSG 49
QY 64 SDVPGLPVNPMLRAASEITLNDGFEVLHDHGLDPLNLRQIGSSV-FRVETOEDGKHIAVG 122
Db 50 SSVTSSGVSTATISGSSVTSN-GVSIV-TNSEFHTTSGISTATNSESTASSGISIATN 107
QY 123 QRNGVETS--VVLSDQEVARLQSIDPEGKQKVFVFTGGRGGAGHAMVTVASDITEARQIRL 180
Db 108 SESSTTSSGASTATNSE-----SSTPSSGASTVTVNSGSSVTSSGASTATNSSESTVSSRAS 163
QY 181 ELLEPKGTGESKGAGESKGVGELRESNGAENTTETQTSTSSLRSDPKLWLALGTAT 240
Db 164 TATNSESTLSSGASTATN-SDSSTTSSGASTATNSESTTSSG-----ASTAT 211
QY 241 GLIGLAATGIQALALTPEDSPPTTDPDAASAATETATRDQLTKEAFQNPQNKVNIDE 300
Db 212 N-----SESVSSRATATNSESTT---SSGASTATNSESTTSSGAGTATNSESTTS 264
QY 301 LG-----NAIPSGVLKDDVVANIEEQAKAAGEEAKQQAIIENNAQAQKYDEQQAQR 351
Db 265 SGASTATNSDSSTVSSGA---STATNSESTTSSGAST-----ATN 302
QY 352 QEELKVSAGYGL-SGALLGGGIGVAVTAALHRKNQPVQTTTTTTTTTSARTVEN 410
Db 303 SESSTTSSGASTATNSDSSTTSSGAGTATNSESTVSSGISTVTVNSSESTPSSGANTATN 362
QY 411 KPANTPAQGNVDTPGSEDPTMESRRSSMASTSTFFDTSSIGTVQNPVADVKTSS-----464
Db 363 SESSTTSSGANTAT-NSESTVSSGASTATNSESTTSSGVTATNSESTTSSGASTAT 421
QY 465 LHDSQVPTNSNTSVQNMGNDSVVYSTIQHPPRDTTNGARLLGNPSAGIQSTVARLAL 524
Db 422 NSDSTTSSSEASTATNSESTVSSGISTVTVNSSESTTSSGANTATNSGSSVTS-----474
QY 525 SGGLRHDGGLTGGNSAVNTSNPPAPG 553
Db 475 AGSGTAALTGMHTTSHSA-STAVSEAKPG 502

RESULT 7
US-10-176-758-310
; Sequence 310, Application US/10176758
; Publication No. US20030008353A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C104
; CURRENT APPLICATION NUMBER: US/10/176,758
; CURRENT FILING DATE: 2002-06-21
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612

; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-758-310

Query Match
Best Local Similarity 19.2%; Score 180; DB 14; Length 596;
Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;

QY 4 GNLGHNPNVNSIPAPPPLPSQTDGAGRGQLINSTGPLGSRALFTPVNRNSMADSGDNRA 63
Db 13 GLLHLEAATNS-----NETSANTGSSVSSG-----ASTATNSG 49
QY 64 SDVPGLPVNPMLRAASEITLNDGFEVLHDHGLDPLNLRQIGSSV-FRVETOEDGKHIAVG 122
Db 50 SSVTSSGVSTATISGSSVTSN-GVSIV-TNSEFHTTSGISTATNSESTASSGISIATN 107
QY 123 QRNGVETS--VVLSDQEVARLQSIDPEGKQKVFVFTGGRGGAGHAMVTVASDITEARQIRL 180
Db 108 SESSTTSSGASTATNSE-----SSTPSSGASTVTVNSGSSVTSSGASTATNSSESTVSSRAS 163
QY 181 ELLEPKGTGESKGAGESKGVGELRESNGAENTTETQTSTSSLRSDPKLWLALGTAT 240
Db 164 TATNSESTLSSGASTATN-SDSSTTSSGASTATNSESTTSSG-----ASTAT 211
QY 241 GLIGLAATGIQALALTPEDSPPTTDPDAASAATETATRDQLTKEAFQNPQNKVNIDE 300
Db 212 N-----SESVSSRATATNSESTT---SSGASTATNSESTTSSGAGTATNSESTTS 264
QY 301 LG-----NAIPSGVLKDDVVANIEEQAKAAGEEAKQQAIIENNAQAQKYDEQQAQR 351
Db 265 SGASTATNSDSSTVSSGA---STATNSESTTSSGAST-----ATN 302
QY 352 QEELKVSAGYGL-SGALLGGGIGVAVTAALHRKNQPVQTTTTTTTTTSARTVEN 410
Db 303 SESSTTSSGASTATNSDSSTTSSGAGTATNSESTVSSGISTVTVNSSESTPSSGANTATN 362
QY 411 KPANTPAQGNVDTPGSEDPTMESRRSSMASTSTFFDTSSIGTVQNPVADVKTSS-----464
Db 363 SESSTTSSGANTAT-NSESTVSSGASTATNSESTTSSGVTATNSESTTSSGASTAT 421
QY 465 LHDSQVPTNSNTSVQNMGNDSVVYSTIQHPPRDTTNGARLLGNPSAGIQSTVARLAL 524
Db 422 NSDSTTSSSEASTATNSESTVSSGISTVTVNSSESTTSSGANTATNSGSSVTS-----474
QY 525 SGGLRHDGGLTGGNSAVNTSNPPAPG 553
Db 475 AGSGTAALTGMHTTSHSA-STAVSEAKPG 502

RESULT 8
US-10-175-737-310
; Sequence 310, Application US/10175737
; Publication No. US20030013153A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C50
; CURRENT APPLICATION NUMBER: US/10/175,737
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
```

; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-737-310

Query Match 6.3%; Score 180; DB 14; Length 596;
Best Local Similarity 19.2%; Pred. No. 0.0021;
Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;

QY 4 GNLGHNPNVNSIIPAPPPLPSQTDGAGRGQLINSTGLGRALFTPVNRNSMADSGDNRA 63
DB 13 GLLHLEAATNS-----NETSTANTGSSVSSG-----ASTATNSG 49
QY 64 SDVPGLPVNPMLAASEITLNDGPEVLHDHGPDLTLNRQIGSSV-FRVETQBDGKHIAVG 122
DB 50 SSVTSSGVSTATIGSSVTSN-GVSIV-TNSEFHTTSSGISTATNSEFSTASSGSIATN 107
QY 123 QRNGVETS--VVLSDQEVARLQSIDPEGKDFVFTGGRGGAGHAMVTVASDITEARQRL 180
DB 108 SESSTSSGASTATNSE-----SSTPSSGASTVTNNGSSVTSSGASTATNSESTVSSRAS 163
QY 181 ELLPEKGTGESKGAGESKGVGELRESNGAENTTETQTSTSTSSLRSDPKLWLALGTAT 240
DB 164 TATNSESTLSSGASTATN-SDSSTSSGASTATNSESTSSG-----ASTAT 211
QY 241 GLIGLAATGIVOALALTPEPDSPTTDPDAAASATETATRDQLTKEAFQNPQKNWIDE 300
DB 212 N-----SESSTVSSRASTATNSESTT---SSGASTATNSESTTNGAGTATNSESTTS 264
QY 301 LG-----NAIPSGVLKDDVANIIEQAKAAGEAKQAIENNAQAQKYDEQQAQR 351
DB 265 SGASTATNSDSSTVSSGA---STATNSESTSSGAST-----ATN 302
QY 352 QBELKVSAGAGYGL-SGALILGGGIGVAVTAALHRKNQPVETTTTTTTTTTTSARTVEN 410
DB 303 SESSTSSGASTATNSDSSTSSGAGTATNSESTVSSGISTVTNSESTPSSGANTATN 362
QY 411 KPANNTPAQGVNDTPGSEDTHMESRRSSMASTSTFFDTSSIGTVQNPYADVKT----- 464
DB 363 SESSTSSGANTAT-NSESSTVSSGASTATNSESTSSGVTATNSESTSSGASTAT 421
QY 465 LHDQVPTNSNTSVQNMGTDSVVYSTIOHPPTDNGARLLGNPAGIOSTVARIAL 524
DB 422 NSDSTTSSGASTATNSESTVSSGISTVTNSESTSSGANTATNSESTSSVTS----- 474
QY 525 SGGLRHDMDGLTGGNSAVNTSNPPAPG 553
DB 475 AGSGTAALTGMHTTSHSA-STAVSEAKPG 502

RESULT 9

US-10-063-616-100
; Sequence 100, Application US/10063616
; Publication No. US20030013855A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,616
; CURRENT FILING DATE: 2002-05-03
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170

; SEQ ID NO 100
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-063-616-100

Query Match 6.3%; Score 180; DB 14; Length 596;
Best Local Similarity 19.2%; Pred. No. 0.0021;
Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;

QY 4 GNLGHNPNVNSIIPAPPPLPSQTDGAGRGQLINSTGLGRALFTPVNRNSMADSGDNRA 63
DB 13 GLLHLEAATNS-----NETSTANTGSSVSSG-----ASTATNSG 49
QY 64 SDVPGLPVNPMLAASEITLNDGPEVLHDHGPDLTLNRQIGSSV-FRVETQBDGKHIAVG 122
DB 50 SSVTSSGVSTATIGSSVTSN-GVSIV-TNSEFHTTSSGISTATNSEFSTASSGSIATN 107
QY 123 QRNGVETS--VVLSDQEVARLQSIDPEGKDFVFTGGRGGAGHAMVTVASDITEARQRL 180
DB 108 SESSTSSGASTATNSE-----SSTPSSGASTVTNNGSSVTSSGASTATNSESTVSSRAS 163
QY 181 ELLPEKGTGESKGAGESKGVGELRESNGAENTTETQTSTSTSSLRSDPKLWLALGTAT 240
DB 164 TATNSESTLSSGASTATN-SDSSTSSGASTATNSESTSSG-----ASTAT 211
QY 241 GLIGLAATGIVOALALTPEPDSPTTDPDAAASATETATRDQLTKEAFQNPQKNWIDE 300
DB 212 N-----SESSTVSSRASTATNSESTT---SSGASTATNSESTTNGAGTATNSESTTS 264
QY 301 LG-----NAIPSGVLKDDVANIIEQAKAAGEAKQAIENNAQAQKYDEQQAQR 351
DB 265 SGASTATNSDSSTVSSGA---STATNSESTSSGAST-----ATN 302
QY 352 QBELKVSAGAGYGL-SGALILGGGIGVAVTAALHRKNQPVETTTTTTTTTTTSARTVEN 410
DB 303 SESSTSSGASTATNSDSSTSSGAGTATNSESTVSSGISTVTNSESTPSSGANTATN 362
QY 411 KPANNTPAQGVNDTPGSEDTHMESRRSSMASTSTFFDTSSIGTVQNPYADVKT----- 464
DB 363 SESSTSSGANTAT-NSESSTVSSGASTATNSESTSSGVTATNSESTSSGASTAT 421
QY 465 LHDQVPTNSNTSVQNMGTDSVVYSTIOHPPTDNGARLLGNPAGIOSTVARIAL 524
DB 422 NSDSTTSSGASTATNSESTVSSGISTVTNSESTSSGANTATNSESTSSVTS----- 474
QY 525 SGGLRHDMDGLTGGNSAVNTSNPPAPG 553
DB 475 AGSGTAALTGMHTTSHSA-STAVSEAKPG 502

RESULT 10

US-10-174-581-310
; Sequence 310, Application US/10174581
; Publication No. US20030017540A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C41
; CURRENT APPLICATION NUMBER: US/10/174,581
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: 10/052586

1	PRIOR APPLICATION NUMBER: 60/082565
2	PRIOR FILING DATE: 1998-04-21
3	PRIOR APPLICATION NUMBER: 60/082569
4	PRIOR FILING DATE: 1998-04-21
5	PRIOR APPLICATION NUMBER: 60/082704
6	PRIOR FILING DATE: 1998-04-22
7	PRIOR APPLICATION NUMBER: 60/082797
8	PRIOR FILING DATE: 1998-04-22
9	PRIOR APPLICATION NUMBER: 60/083322
10	PRIOR FILING DATE: 1998-04-28
11	PRIOR APPLICATION NUMBER: 60/083495
12	PRIOR FILING DATE: 1998-04-29
13	PRIOR APPLICATION NUMBER: 60/083496
14	PRIOR FILING DATE: 1998-04-29
15	PRIOR APPLICATION NUMBER: 60/083499
16	PRIOR FILING DATE: 1998-04-29
17	PRIOR APPLICATION NUMBER: 60/083559
18	PRIOR FILING DATE: 1998-04-29
19	PRIOR APPLICATION NUMBER: 60/084366
20	PRIOR FILING DATE: 1998-05-05
21	PRIOR APPLICATION NUMBER: 60/084414
22	PRIOR FILING DATE: 1998-05-06
23	PRIOR APPLICATION NUMBER: 60/084639
24	PRIOR FILING DATE: 1998-05-07
25	PRIOR APPLICATION NUMBER: 60/085573
26	PRIOR FILING DATE: 1998-05-15
27	PRIOR APPLICATION NUMBER: 60/085579
28	PRIOR FILING DATE: 1998-05-15
29	PRIOR APPLICATION NUMBER: 60/085580
30	PRIOR FILING DATE: 1998-05-15
31	PRIOR APPLICATION NUMBER: 60/085582
32	PRIOR FILING DATE: 1998-05-15
33	PRIOR APPLICATION NUMBER: 60/085700
34	PRIOR FILING DATE: 1998-05-15
35	PRIOR APPLICATION NUMBER: 60/086023
36	PRIOR FILING DATE: 1998-05-18
37	PRIOR APPLICATION NUMBER: 60/086392
38	PRIOR FILING DATE: 1998-05-22
39	PRIOR APPLICATION NUMBER: 60/086486
40	PRIOR FILING DATE: 1998-05-22
41	PRIOR APPLICATION NUMBER: 60/087098
42	PRIOR FILING DATE: 1998-05-28
43	PRIOR APPLICATION NUMBER: 60/087208
44	PRIOR FILING DATE: 1998-05-28
45	PRIOR APPLICATION NUMBER: 60/087609
46	PRIOR FILING DATE: 1998-06-02
47	PRIOR APPLICATION NUMBER: 60/087759
48	PRIOR FILING DATE: 1998-06-02
49	PRIOR APPLICATION NUMBER: 60/088025
50	PRIOR FILING DATE: 1998-06-04
51	PRIOR APPLICATION NUMBER: 60/088028
52	PRIOR FILING DATE: 1998-06-04
53	PRIOR APPLICATION NUMBER: 60/088029
54	PRIOR FILING DATE: 1998-06-04
55	PRIOR APPLICATION NUMBER: 60/088033
56	PRIOR FILING DATE: 1998-06-04
57	PRIOR APPLICATION NUMBER: 60/088167
58	PRIOR FILING DATE: 1998-06-05
59	PRIOR APPLICATION NUMBER: 60/088202
60	PRIOR FILING DATE: 1998-06-05
61	PRIOR APPLICATION NUMBER: 60/088212
62	PRIOR FILING DATE: 1998-06-05
63	PRIOR APPLICATION NUMBER: 60/088217
64	PRIOR FILING DATE: 1998-06-05
65	PRIOR APPLICATION NUMBER: 60/088326
66	PRIOR FILING DATE: 1998-06-04
67	PRIOR APPLICATION NUMBER: 60/088655

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Db 422 NSDSTTSSEASTATNSSESTVSSGISTVTNSSESTTSGGANTATNSGSSVTS----- 474
QY 525 SGGLRHDGGLTGGNSAVNTSNPPAPG 553
Db 475 AGSGTAALTGMHTTSHSA-STAVSEAKPG 502

RESULT 12

US-10-176-749-310
; Sequence 310, Application US/10176749
; Publication No. US20030017542A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C76
; CURRENT APPLICATION NUMBER: US/10/176,749
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-749-310

Query Match 6.3%; Score 180; DB 14; Length 596;
Best Local Similarity 19.2%; Pred. No. 0.0021;
Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;
QY 4 GNLGHNPNVNSIPPAPLPSTQDAGGGRGQLINSTGPLGRALFTPVNRNSMADSGDNRA 63
Db 13 GLLHLEAATNS-----NETSTANTSGSVISSG-----ASTATNSG 49
QY 64 SDVPGLPVNPRLAASEITLNDGFEVLHDHGPDLTLNRQIGSSV-FRVETQEDGKHIAVG 122
Db 50 SSVTSSGVSTATISGSSVTSN-GVSIV-TNSEFHTTSSGISTATNSSEFSTASSGISIATN 107
QY 123 QRNGVETS--VVLSDQEVARLQSIDPEGKDKFVFTGGRGGAGHAMVTVASDITEARQRIL 180
Db 108 SESSTTSSGASTATNSE-----SSTPSSGASTVTNSGSSVTSSGASTATNSSESTVSSRAS 163
QY 181 ELLPEKGTGESKAGESKGVGELRESNGAENTTETQSTSTSSLRSDPKLWLALGTVAT 240
Db 164 TATNSSESTLSSGASTATN-SDSSTTSSGASTATNSSESTTSSG-----ASTAT 211
QY 241 GLIGLAATGIVOALALITPEPDSPTTDPDAASATETATRDOLTKEAFQNPDPNQKWNIDE 300
Db 212 N-----SESSTVSSRASTATNSSESTT-----SSGASTATNSSESTTNGAGTATNSSESTTS 264
QY 301 LG-----NAIPSGVLKDDVVANIEEQAKAAGEAKQQAIAENNAQAQKYDQQAQR 351
Db 265 SGASTATNSDSTVSSGA-----STATNSSESTTSSGAST-----ATN 302
QY 352 QEELKVSAGAGYGL-SGALILGGGIGVAVTAALHRKNQFVEQTITTTTTTTTTTARTVEN 410
Db 303 SESSTTSSGASTATNSDSTTSSGAGTATNSSESTVSSGISTVTNSSESTPSSGANTATN 362
QY 411 KPANNTPAQGVNDTPGSEDTWESRRSSMASSTSTFTDSSIGTGVONPYADVKTSS----- 464
Db 363 SESSTTSSGANTAT-NSSESTVSSGASTATNSSESTTSSGVTATNSSESTTSSGASTAT 421

QY 465 LHDSQVPTNSNTSVQNMGNNTDSVVYSTIQHPRPDRTDNGARLLGNPSGAGIQSTVARLAL 524
Db 422 NSDSTTSSEASTATNSSESTVSSGISTVTNSSESTTSGGANTATNSGSSVTS----- 474
QY 525 SGGLRHDGGLTGGNSAVNTSNPPAPG 553
Db 475 AGSGTAALTGMHTTSHSA-STAVSEAKPG 502

RESULT 13

US-10-176-914-310
; Sequence 310, Application US/10176914
; Publication No. US20030017543A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C83
; CURRENT APPLICATION NUMBER: US/10/176,914
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-914-310

Query Match 6.3%; Score 180; DB 14; Length 596;
Best Local Similarity 19.2%; Pred. No. 0.0021;
Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;
QY 4 GNLGHNPNVNSIPPAPLPSTQDAGGGRGQLINSTGPLGRALFTPVNRNSMADSGDNRA 63
Db 13 GLLHLEAATNS-----NETSTANTSGSVISSG-----ASTATNSG 49
QY 64 SDVPGLPVNPRLAASEITLNDGFEVLHDHGPDLTLNRQIGSSV-FRVETQEDGKHIAVG 122
Db 50 SSVTSSGVSTATISGSSVTSN-GVSIV-TNSEFHTTSSGISTATNSSEFSTASSGISIATN 107
QY 123 QRNGVETS--VVLSDQEVARLQSIDPEGKDKFVFTGGRGGAGHAMVTVASDITEARQRIL 180
Db 108 SESSTTSSGASTATNSE-----SSTPSSGASTVTNSGSSVTSSGASTATNSSESTVSSRAS 163
QY 181 ELLPEKGTGESKAGESKGVGELRESNGAENTTETQSTSTSSLRSDPKLWLALGTVAT 240
Db 164 TATNSSESTLSSGASTATN-SDSSTTSSGASTATNSSESTTSSG-----ASTAT 211
QY 241 GLIGLAATGIVOALALITPEPDSPTTDPDAASATETATRDOLTKEAFQNPDPNQKWNIDE 300
Db 212 N-----SESSTVSSRASTATNSSESTT-----SSGASTATNSSESTTNGAGTATNSSESTTS 264
QY 301 LG-----NAIPSGVLKDDVVANIEEQAKAAGEAKQQAIAENNAQAQKYDQQAQR 351
Db 265 SGASTATNSDSTVSSGA-----STATNSSESTTSSGAST-----ATN 302
QY 352 QEELKVSAGAGYGL-SGALILGGGIGVAVTAALHRKNQFVEQTITTTTTTTTTTARTVEN 410
Db 303 SESSTTSSGASTATNSDSTTSSGAGTATNSSESTVSSGISTVTNSSESTPSSGANTATN 362
QY 411 KPANNTPAQGVNDTPGSEDTWESRRSSMASSTSTFTDSSIGTGVONPYADVKTSS----- 464
Db 363 SESSTTSSGANTAT-NSSESTVSSGASTATNSSESTTSSGVTATNSSESTTSSGASTAT 421

[illegible]

```

RESULT 14
US-10-176-915-310
; Sequence 310, Application US/10176915
; Publication No. US20030017544A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C110
; CURRENT APPLICATION NUMBER: US/10/176,915
; CURRENT FILING DATE: 2002-06-21
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 310
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-915-310

```

Query Match	6.3%;	Score 180;	DB 14;	Length 596;
Best Local Similarity	19.2%;	Pred. No. 0.0021;		
Matches	109;	Conservative	90;	Mismatches 272; Indels 98; Gaps 19;
QY	4	GNLGHNNVANSIPPAPPLPSQTDGAGRGQLINSTGLGSRALFTVPRNSMADSGDRA	63	
Db	13	GLLHLLEAATNS-----NETSTANTGSSVISSG-----ASTATNSG	49	
QY	64	SDVPGFLVPMRLAASEITLNDGFEVLHDHGPDLTLNRQIGSSV-FRVETQEDCKHIAVG	122	
Db	50	SSVTSSGCVSTATISGSSVTSN-GVSIIV-TNSEFHTTSSGISTATNSSEFSTASSGISIATN	107	
QY	123	QRNGVETS--VVLSDQYARLQSIDPGKOKFVETGGRGGAGHAMVTVASDITEARQIL	180	
Db	108	SESSTTSSGASTATNSE-----SSTPSSGASVTNMSGSVTSSGASTATNSSESSTVSSRAS	163	
QY	181	ELLEPKGTGSKGAGESKGVCELRESNGSAGENTTETOTSTSTSLRSDPKLWLALGTVAT	240	
Db	164	TATNSSESTLSSGASTATN-SDSSTTSSGASTATNSSESSTTSSG-----ASTAT	211	
QY	241	GLICLAATGIVQALATPEPDSPTTTPDPAASATETATRDQLTKEAFQNPNDOKNVIDE	300	
Db	212	N-----SESSTVSSRASATNSSESSTT---SSGASTATNSSESRITNSGAGTATNSSESSTTS	264	
QY	301	LG-----NAIPSGVLKDDVVANIEEQAKAAGEAKQQAIEENNAQAOKKYDEQQAQR	351	
Db	265	SGASTATNSDSSTVSSGA-----STATNSSESSTTSSGAST-----ATN	302	
QY	352	QEBELKVSSGAGYGL-SCALILGGGIVGAVTAALHRKNQPVQETTTTTTTTTTTSARTVEN	410	
Db	303	SESSTTSSGASTATNSDSSTTSSGAGTATNSSESTVSSGISTVTNSESSTPSSGANTATN	362	
QY	411	KPANNTPAQGNVTPGSEDWTMESRRRSMASSTSTFTFTSSIGTVQNFPYADVKTIS-----	464	

Db	363	SESSTSSGANTAT-NSESSTVSGSGATATNSESSTSSGVSATNSESSTSSGASTAT	421
Qy	465	LHDSQVTSNTSVQNMGNITDSVVYSTIQHPDRDITDNGARLLGNPSAGIOSTVARLAL	524
Db	422	NSDSSTTSSGASTATNSESSTVSGISTVTNSESSTSSGANTATNCGSVTS-----	474
Qy	525	SGGLRHDMGGITGCGNSGAVNTSNPPAPG	553
Db	475	AGSGTAALTGMMHTTSHSA-STAVSEAKPG	502

RESULT 15	US-10-063	Publica	GENERAL	APPLIC	APPLIC	APPLIC	APPLIC	APPLIC	APPLIC	APPLIC	FILE	CURREN	CURREN	Prin	NUMBER	SEQ ID	LENGTH	TYPE:	US-10-063
		Publication	GENERAL	APPLIC	APPLIC	APPLIC	APPLIC	APPLIC	APPLIC	APPLIC	FILE	CURREN	CURREN	Prin	NUMBER	SEQ ID	LENGTH	TYPE:	US-10-063

Query Match .	6.3%;	Score 180;	DB 14;	Length 596;
Best Local Similarity	19.2%;	Pred. No. 0.0021;		
Matches 109;	Conservative	90;	Mismatches 272;	Indels 98;
Gaps 19				
QY	4	GNLGHNPVNVNSPPAPPLPSQTDGAGRGQLINSTPLGSRALFPTVRNSMADSGDNRA	63	
Db	13	GLLIHLLEAATNS-----NETSTANTGSSVISSG-----ASTATNSG	49	
QY	64	SDVPGPLVPNWRLLAASEITLNDGPEVLHDHCPDLTLNRQIGSSV-FRVETQEDGKHIAVG	122	
Db	50	SSVTSGSVSTATISGSSVTGN-GVSIW-TNSEFFTTSGISTATNSSEFSTASSGSIATN	107	
QY	123	QRNGVETS--VVLSDQRYARLQSIDPBGKQKVFETGRRGGAGHAMVTVASDITEARQIL	180	
Db	108	SESSTTSSGASTATNSE-----SSTPSSGASTVTNNGSSVTSSGASTATNSSESTVSSRAS	163	
QY	181	ELLEPKGTGSKGAGESKGVGELRESNGAENTTETQTSTSTSLRSDPKLWLALGTVAT	240	
Db	164	TATNSSESTLSSGASTATN-SDSSTTSSGASTATNSSESTTSSG-----ASTAT	211	
QY	241	GLIGLAATGIVQALALTPEDSPSTTTDPDAAAASATETATRDQLTKEAFPONDPKNKWNIDE	300	
Db	212	N-----SESSTVSSRASTATNSSESTT---SSGASTATNSSESTTSSNGAGTATNSSESTTS	264	
QY	301	LG-----NAIPSGVLKDDVVANIEEQAAGEAKQOAIENNAQAOKYDEQQAQR	351	
Db	265	SGASTATNSDSSTVSSGA---STATNSSESTTSSGAST-----ATN	302	
QY	352	QEBLKVSSSGAGYGL-SCALILGGIGIVAVTAALHRKNQPVQETTTTTTTTTTISARTVEN	410	
Db	303	SESSTTSSGASTATNSDSSTTSSGAGTATNSSESTVSSGISTVTNNSSESTPSSGANTATN	362	
QY	411	KPANNTPAQGNVDTPGSEDTWERRRRSMASTSTSTFFDTSSIGTVQNPNYADVKTS-----	464	
Db	363	SESSTTSSGANTAT-NSESTVSSGASTATNSSESTTSSGUSTATNSSESTTSSGASTAT	421	

Qy	465	LHDSQVTSNSNTSVQNMGNLTDVVYVSTIQHPDRDTDNGARLLGNPAGIQSTYARLAL	524
Db	422	NSDSSTTSSEASTATNSESTVSSGISVTNSESSTSSGANTATNSGSSVTS-----	474
Qy	525	SGGLRHDMDGGLTGGNSAVNTSNNPPAPG	553
Db	475	AGSGTAALTGMHTTSHSA-STAVSEAKPG	502

Search completed: May 13, 2005, 12:00:29
Job time : 138 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 13, 2005, 11:37:06 ; Search time 43 Seconds
(without alignments)
968.702 Million cell updates/sec

Title: US-09-189-415B-11
Perfect score: 2840
Sequence: 1 MPIGNLGNPNVNSIPPAP.....SNSAVNTSNPPAPGSHRFV 558

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pap:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pap:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pap:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pap:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pap:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	172	6.1	2137	3	US-09-134-001C-4463
2	172	6.1	2736	4	US-09-252-991A-30227
3	170	6.0	529	4	US-09-248-796A-16703
4	166	5.8	2870	4	US-09-479-467A-15
5	166	5.8	3178	4	US-09-479-467A-15
6	163.5	5.8	821	4	US-09-556-877-195
7	163.5	5.8	821	4	US-09-620-412C-195
8	163.5	5.8	821	4	US-09-598-419-195
9	163.5	5.8	1776	4	US-09-556-877-179
10	163.5	5.8	1776	4	US-09-620-412C-179
11	163.5	5.8	1776	4	US-09-598-419-179
12	159.5	5.6	441	4	US-09-248-796A-15936
13	158	5.6	556	4	US-09-248-796A-22338
14	157.5	5.5	827	4	US-09-248-796A-17307
15	152.5	5.4	2045	4	US-09-949-016-10491
16	152	5.4	2314	3	US-09-268-347-49
17	151	5.3	1306	4	US-09-538-092-330
18	149.5	5.3	785	4	US-09-538-092-872
19	148	5.2	2035	1	US-08-046-585-5
20	148	5.2	2035	1	US-08-393-703-5
21	148	5.2	2035	5	PCT-US93-11721-5
22	148	5.2	2142	4	US-09-540-236-3459
23	147	5.2	525	4	US-09-107-532A-5095
24	147	5.2	985	5	PCT-US96-03916-6
25	147	5.2	985	5	PCT-US96-03916-66
26	146.5	5.2	1095	4	US-09-206-942-69
27	146.5	5.2	1536	1	US-08-038-682-2

28	146.5	5.2	1536	1	US-08-302-832-2	Sequence 2, Appli
29	146.5	5.2	1536	2	US-08-530-138-2	Sequence 2, Appli
30	146.5	5.2	1536	2	US-08-463-880-2	Sequence 2, Appli
31	146.5	5.2	1536	2	US-08-728-470-2	Sequence 2, Appli
32	146.5	5.2	1536	2	US-08-617-697-2	Sequence 2, Appli
33	146.5	5.2	1536	3	US-08-719-641-2	Sequence 2, Appli
34	146.5	5.2	1536	4	US-09-206-942-67	Sequence 67, Appli
35	145.5	5.1	3241	4	US-09-841-786-1	Sequence 1, Appli
36	145	5.1	461	1	US-08-186-222-2	Sequence 2, Appli
37	145	5.1	1938	4	US-09-949-016-6609	Sequence 6609, Ap
38	144.5	5.1	2843	1	US-07-741-940-2	Sequence 2, Appli
39	144.5	5.1	2843	1	US-08-289-548A-2	Sequence 2, Appli
40	144.5	5.1	2843	1	US-08-453-654-2	Sequence 2, Appli
41	144.5	5.1	2843	2	US-08-370-235A-2	Sequence 2, Appli
42	144.5	5.1	2843	4	US-08-443-731-2	Sequence 2, Appli
43	144	5.1	1529	2	US-08-728-470-10	Sequence 10, Appli
44	144	5.1	1529	3	US-08-719-641-10	Sequence 10, Appli
45	144	5.1	1600	2	US-08-617-697-10	Sequence 10, Appli

ALIGNMENTS

RESULT 1
US-09-134-001C-4463
; Sequence 4463, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4463
; LENGTH: 2137
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4463

Query Match		6.1%;	Score 172;	DB 3;	Length 2137;
Best Local Similarity		19.3%;	Pred. No. 0.00031;		
Matches 117;		Conservative 98;	Mismatches 281;	Indels 110;	Gaps 20;
QY	36	INSTGPIGSRAL----	FTFVNSMADSGDNRASDPGLPVPNPMRLAASEIT-LNDGFEVL	90	
DB	776	TESTDASGNKTTTKINYEVTRNSASDSTSTSI-----VNSVSTSI	NSSTLSDSVKAS	828	
QY	91	HDHGLDLTLNRIGQSSVFRVETOEDGKHIAVGORGVETSVVLSD--	OBYAR-----	140	
DB	829	QSLSTKSLSSLSASTNSNSTSIQASASASTSKQLSASASTSDSASAKS	SESTSKS	888	
QY	141	---LQSIDPEGKDKFVFTGGRGAGHAMV-----TVASDIT-----	BARQRI	179	
DB	899	TSLSESTSVSDSASVSTSESASTSVSGSTSTSI	SDSTSTSDGSAIKASEASTS	948	
QY	180	LELLPKGTGSGKAGESKGVGELRENSGANTTETQTSTSTSLRDPKWLALG	TVTA	239	
DB	949	KLLSESVSTSDSASTSTSVSDNSASTSKSTSTSVSDSTSTSDS----	ASTSTST	1004	
QY	240	TGLIGLAATGIVQALALTPEPDPSTTTPDAAASATATATRDQLT-----	KEA	287	
DB	1005	REESDASTSTSESTSTSVSDSTSTSDSASMSASENSKSTSLSESTSTSLSG	STST	1064	
QY	288	FONPNQKVNIDELGNALPISGVLKDDVVANIEQAKAA--GEEAKQQAIIENNAQ	AKK----	343	
DB	1065	ASTSDASTSTSESDSTSTSLSESTSTSLSGSTASSTSDSASTSTSESDST	SESTSL	1124	

QY 344 -----YDEQAKQBEELKVSSGAGYGLSGALILGGIGVAVTAALHRKNQPVQETTT- 395
Db 1125 EBLSTSVSDSTASSTSEASSTSESESNASTSLGSLSTISDSSTSTSDSASTSTS 1184
QY 396 -----TTTTTTTARTVENKANNTPAQGNVDPGSEDIMESRSSMASTSTSTFFDTS 449
Db 1185 ESESSTSLSESTSLSDSTSTSTSESASTSTSESASTSTSES-TSLSESTSTSVSDST 1243
QY 450 SIGTVQNPYADVKTSLHDSQ-VPTGNS---NTSVQNMGT---DSVYSTTQHPPRDTTD 502
Db 1244 SASTSDS---ASTSTSVSDSEASSTSLSESLSTSVSDSTSTSDSASTSTSE---SDSTS 1298
QY 503 NGARLLGNPSAGI-----QSTYARLALSGGURHDMGGLTGSNS-A 542
Db 1299 ESTLSSESTSVSDSTGASTSDSASTSTSESESDSASTSLSGSTSTSLSDSTSTSDS 1358
QY 543 VNTSNN 548
Db 1359 ASTSTS 1364

RESULT 2
US-09-252-991A-30227
; Sequence 30227, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30227
; LENGTH: 2736
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30227

Query Match 6.1%; Score 172; DB 4; Length 2736;
Best Local Similarity 23.9%; Pred. No. 0.00045;
Matches 148; Conservative 60; Mismatches 242; Indels 170; Gaps 33;
QY 2 PIGNLGHNPVNNSTIPPAPLPSQTDGAGRGQLNSTGPLGSRALFTFVRNSMADSDN 61
Db 723 PIGQVTADSGSNWSFTPTPLPN-----GTVVNATATDAS-----GNTSAGSSVT 767
QY 62 RASDVPGPLPV-NPMRLAASEITLNDGFVLHDHGFPLDLNRQIGSSVFRVETQEDGKHIA 120
Db 768 VDSVAPATPVINP-----SNGTTLSTGTAEP-----GSSV-----TLTDGNGNP 805
QY 121 VGORN-----GVETSVVLSDOEVARLQSIDPEGDKFVFTGGRGGA---GHAMVTVAS 170
Db 806 IQQVTTADSGSNWSFTPSPLADGTVVNATATDAGN-----TSGGSGTVDGVAPTFTTV 860
QY 171 DITEARQRIELLEPKGTGESKGA-----GESKGVGELRENSNGAENTTETQTSTSSSL 225
Db 861 NLSNGSSL-----SGTAEPGSTVILTGDNGNPFAEVTADGSG--NWYTT----- 902
QY 226 RSDPKLWALGTAVTGLTGLAATGIVQALALTPEPDSPT-----TTDPPAA 271
Db 903 ---PSTPIANGTVNVNVAQDAAGNSPGASVTVDSPAAPVNPVNSGTTLSGTAEPGAT 959
QY 272 ASATETATRDQLTKEAFQNPQKNIDELGN--AIPSGVLKDDVVANIEQAKAAGEA 329
Db 960 VSLT-----DNGNPIQV--VTADSGSNWSFTPGTPLANGTVNVNATASDPTGNTSA 1008
QY 330 -KQQAIIENNAQKQYDEQAKQBEELKVSSGAGYGLSGALILGGIGVAVTAALHRKNQ 388

Db 1009 PASTTVDSVAPAAVVPNSNG-----VVISGTAEPGATVTLTDGSG-----N 1050
QY 389 PVEQTTT-----TTTTTTTARTVEN-----KPAANTPAQGN--VD--TPGESDPTMESR 434
Db 1051 PIGQVTADSGSNWSFTSTPLADGTVVNATATDAGTGGGSGTTVDALIAPATPVNLSN 1110
QY 435 RSSMASTS---STPFDTSSIGTVQNPYADVKT--SLHDSQVP-TSNNGTSVQNMGTDSV 488
Db 1111 GSSLSGTAEPGSTVILTGDNG--NPIAEVTADSGSNWYTPSTPIANGTVNVNVAEDAA 1167
QY 489 VYSTIQHPPRDTT-----DNGARLLGNPSAGIQTSTYARLALSGGLRHDGGL 535
Db 1168 GNSS---PPATVTVDSAPPAPVNPNSGVVISGTAEGATVT---LTDAGG--NPICQV 1219
QY 536 TG-GSNNAVNTSNNPPAPGS 554
Db 1220 TADSGSNWSFTPGTTPPANGT 1239

RESULT 3
US-09-248-796A-16703
; Sequence 16703, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 16703
; LENGTH: 529
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-16703

Query Match 6.0%; Score 170; DB 4; Length 529;
Best Local Similarity 23.0%; Pred. No. 5.7e-05;
Matches 85; Conservative 47; Mismatches 150; Indels 88; Gaps 15;
QY 205 ESNSGAENTTQTSTSTSSLSRSDPKLWALGTAVTGLGLAATGIVQALALTPEPDSPT 264
Db 208 QSTSGKTTTTPVVTATSTAGNDN---TASTTDATG-----KTLTTVTSNDNTT 255
QY 265 TTDPAASAATETATRDQLTKEAFQNPQKNIDELGNALPISGVLKDDVVANIEQAK- 323
Db 256 STGDDSTTASTGNDNTSTTTATVTK-----NIDSTTNA-TDNTSKATTTGNDNTDST 309
QY 324 -AAGEERAKQAIIENNAQ---AQKQYDEQAKQBEELKVSSGAGYGLSGALILGGIGVAV 379
Db 310 TASGTDSTTTTGDNNSESTTASTGNDD-----NDSTTVTTGDDY-----TTV 351
QY 380 TAALHRKNQVQETTTTTTTTTTTSARTVENKPAANTPAQGNVD---TPGESDPTMESRR 435
Db 352 TT-----DNDNTASIVTTGNSNTTTLTDATGKDSSTTTTNDGHDESTTVTTGDDDTATTKD 407
QY 436 SSMASTSTSTFPDTSIGTVQNPYADVKTSLHDSQVPTSNSTSVQNMGTDSVVYST--- 492
Db 408 NNTASTTTKGDNDTATVNH-----SISTVTTTGDNDTATTGNDTDTATGND 460
QY 493 -----IQHPPRDTTNGARLLGNPSAGIQTSTYARLALSGGLRHDGGLTGSNSA 542
Db 461 STTATTGNHSIESTPVTGDDNTATTGNDTAA--STVS-----TGNDNSA 503
QY 543 VNTSNNPAP 552
Db 504 SKTV--PPTP 511

```

RESULT 4
US-09-479-467A-15
; Sequence 15, Application US/09479467A
; Patent No. 6723557
; GENERAL INFORMATION:
; APPLICANT: Sternberg, Paul W.
; APPLICANT: Barr, Maureen M.
; TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE GENE HOMOLOGS REQUIRED FOR MALE MATING
; TITLE OF INVENTION: BEHAVIOR IN NEMATODES AND ASSAYS BASED THEREON
; FILE REFERENCE: 18021-2901B
; CURRENT APPLICATION NUMBER: US/09/479,467A
; CURRENT FILING DATE: 2000-01-06
; PRIOR APPLICATION NUMBER: 60/115,127
; PRIOR FILING DATE: 1999-01-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 2870
; TYPE: PR
; ORGANISM: C. Elegans Lov-1 sy582 deletion protein
US-09-479-467A-15

```

Query Match	5.8%; Score 166; DB 4; Length 2870;
Best Local Similarity	20.3%; Pred. No. 0.0014;
Matches	108; Conservative 89; Mismatches 223; Indels 112; Gaps 19;
QY	68 GLPNPRLAAASEITLND-----GFEVLHDHGLPLDLNROIGSSVPRVETQEDGKHI 119
DB	137 GLFLNSTWITLNEVNDDEISIAVEAKIEVCYODG---IDRCDSGLWV---LQVGNEW 189
QY	120 A-VGORGNGVETSVVLSQDEYARLQSDIPEGKDKFVFTGGRGGAGHAMVTVTASDITEARQ 178
DB	190 ALLGYREKCESGBI--NEEVARMCKRPYRSEK-----STAISDSQGV 230
QY	179 ILELLPKGTGSEKAGESKGVGELRESNGANTTETQTSTSTSLRSDPKLWLALGTV 238
DB	231 YYDQGVLLGVYAKQFSMRSTSGSPTLRMKRDAGDNTCDYTIESTSTSTTTPTTTTSTV 290
QY	239 -ATGLIGLAATGIVQALALTPEDSPITT-----DPDAAASATETATRDOLTKAEFQ 289
DB	291 TSTTTTPTSTVTTTAMSTSTSTPSTSTTTTSTSTSTSTSTSTSTSTSTSTSTSTSTST 350
QY	290 NPDNQKVNIDELGNAIPSGVLKD--DVYANIEBQAKAAG--EAKQQAENNAQAKKYDE 346
DB	351 SPSS-----TTLSTSIPTTTTPTSTSLSLPDNAICSYLDETTSTFTTTMLTSTTTE 405
QY	347 QQAKROBELKVSAGYGLSGALILGGGIGVAVTAALHRKNQVPVEOTTTTTTTTTTTT--S 404
DB	406 EPSTSTTTTEVTS-----TSTVTTTTEPTTTLTSTASTSTEPS 445
QY	405 ARTVENKPA-----NNTPAQGVNDTPGSEDTWESRBSNMASTSSIFDFTSSIGTVQ 455
DB	446 TSTVTTSPSTSPVTSTVTSSSSTTTTPTSTSTSTSTSTSTSTSTSTSTSTSTSTSTST 505
QY	456 NPYADVKTSLHDSQVPTNSGNTSVQNMGNNTDSVVYSTIQHPPRDITDNGARLLGNP---- 511
DB	506 S-----SSTPSSTASSVSSTASSQSSSTSTQSSSTTKSETTSSDGT---NDDPYF 555
QY	512 -----SAGIOSTYARLALSGGLRHDHMGGLTGGNSAVNTSNPPAPGSHRFV 558
DB	556 VERATTTFYDOSTVNLTLNSGL-----GIIGYOTISICTS-----PTSSNRY 597

RESULT 5
US-09-479-467A-4
; Sequence 4, Application US/09479467A
; Patent No. 6723557
; GENERAL INFORMATION:
; APPLICANT: Sternberg, Paul W.
; APPLICANT: Barr, Maureen M.
; TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE
; TITLE OF INVENTION: BEHAVIOR IN NEMATODES AND ASSAYS BASED THEREON
; TITLE OF INVENTION: GENE HOMOLOGS REQUIRED FOR MALE MATING

```

; FILE REFERENCE: 18021-2901B
; CURRENT APPLICATION NUMBER: US/09/479,467A
; CURRENT FILING DATE: 2000-01-06
; PRIOR APPLICATION NUMBER: 60/115,127
; PRIOR FILING DATE: 1999-01-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 3178
; TYPE: PRT
; ORGANISM: C. Elegans Lov-1 protein
US-09-479-467A-4

Query Match      5.8%; Score 166; DB 4; Length 3178;
Best Local Similarity 20.3%; Pred. No. 0.0017;
Matches 108; Conservative 89; Mismatches 223; Indels 112; Gaps 19;

Qy 68 GLPNPMLAASEITLND-----GFEVLHDHGLPLDLNROIGSSVFRVETQEDCKHI 119
Db 137 GLFLNTWITFLNEVNDDEISIAVEAKYEVICYDDG-----IDRCGSLMW---LQVGNEM 189

Qy 120 A-VQORNGVTSVVLSDOEYARLQSDPEKQKDFVTGGRGGAGHAMVTVASDITEARQR 178
Db 190 ALLGYREKCESGHI--NEEYARMCKPYSEK-----STAISDSQGV 230

Qy 179 ILELLPEKGTGSKGAGESKGVGELRESNGAENTTETQTSTSTSLRSDPKMLALGTV 238
Db 231 YDGQVLKGVRAQKQFMRSTSGSPTLRRMKRDAGDNTCDYTIESTSTSTTTPTTTTSTV 290

Qy 239 -ATGLGLAATGIVQALALTPEDSPFTT-----DPAAASATATARDQLTKEAFQ 289
Db 291 TSTTTPSTSTVTTTAMSTSTSTPSTSTTSTTSTSTSTSTSTSTSTSTSTSTSTSTST 350

Qy 290 NPDNQKNIDELGNATPSGVKLD--DVVANIEQAKAAG--EAKQQAENNAQAQKKYDE 346
Db 351 SPSS-----TTLSTSPITTTTPEITSTLSLPDNAICSYLDETTTSTTTTMTLTSTTE 405

Qy 347 QAKROEELKVSAGYGLSGALILGGGIGVAVTAALHRKNQPVQETTTTTTTTTT--S 404
Db 406 EPSTSTTTTTEVTS-----TSSTVTTTTEPTTLTSTASTSTSTSTSTSTSTSTSTST 445

Qy 405 ARTVENKPA-----NNTPAQGVNDTPGSEDTHESRRSSWASTSTSTFTDSSIGTVQ 455
Db 446 TSTVTTSPSTSPVTSTVTSSTSSSTTTTPTSTESTSTSPSSVTSTSTTAPSTSTTGPSS 505

Qy 456 NPYADVKTSLHDQVPTNSNTSVQNMGNDSVVYSTIQHPPRDITDNGARLLGNP--- 511
Db 506 S-----SSTPSTALSSVSSTASSQTSSTQSSSTTKSETTTSSDGT----NPDPFY 555

Qy 512 -----SAGIOSTYARLALSGLRHDMGGLTGGGNSAVNTNNPPAGSHRFV 558
Db 556 VEXATTTFYDSTSNLTNSGL-----GIIGYQTSICTS-----PTSSNNV 597

RESULT 6
US-09-556-877-195
; Sequence 195, Application US/09556877
; Patent No. 6432916
; GENERAL INFORMATION:
; APPLICANT: Probst, Peter
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir
; APPLICANT: Fling, Steve
; APPLICANT: Maisonneuve, Jeff
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C5
; CURRENT APPLICATION NUMBER: US/09/556,877
; CURRENT FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 195
; LENGTH: 821

```

```
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-556-877-195

Query Match
  5.8%; Score 163.5; DB 4; Length 821;
Best Local Similarity 21.9%; Pred. No. 0.00035;
Matches 127; Conservative 67; Mismatches 216; Indels 171; Gaps 27;

Qy 79 SEITLNDGFEVLHDHGLDPLTLNRQIGSSVFRVETQED--GK----- 117
Db 276 TEQTKSNG-----NQDGSSETKDTQVSESPESPDPDDVLGKGGIYTEKSLTITGITGTI 331
Qy 118 ----HIAVQRNGVETSVVLSDQEVARLQSDIDPEGKDFVFTGGRGGAGHAMVTVASDIT 173
Db 332 DFVSNIAATDSGAGVFTKENLSCTNTNSLQFLKN-----SAGQHGGG-AVVTQMSVT 382
Qy 174 EARQRIELLEPKGTGE-----SKGAGESKGVGELRESNGAENTTETOTSTSTSLR 226
Db 383 NTTSE--SITTPPLVGEVIFSENTAKHGCGGICTNKLSLN-----LKTVTLTKNSAKE 434
Qy 227 SDPKLWLALGTVATGLIGLAATGIVQALALTPEDSPPTTDP----- 268
Db 435 SG-----GAFTDLASIPITD-----TPESSTPSSSPASTPEVVASAKINRFPAS 480
Qy 269 ---DAAASATETATRDQL--TKEAFQNPDPNOKVNIDELGNAPSGVLKDDVVANIEEQA 323
Db 481 TAEPAAPSLTE-AESDQTDQETSDTNSDID-VSIENILN-----VAIQNTSAK 528
Qy 324 AA----GEEAKQQAENNAQAQKYDEQAQKQBELKVSSGAGYGLSGALLILG-----GG 374
Db 529 KGGAIYGGKAKLSRINN-----LELSGNSQDVGGGLCLITESVEFDA 570
Qy 375 IGVAVT-----AALHRKNQVPEQTTTTTTTTTTTSARTVEN---KPAANTPAQG-- 420
Db 571 IGSLLSHVNSAAKEGGVTHSKVTLSNLKSTFTPADNTVKAIVESTPEAPEIIPVEGEE 630
Qy 421 -----NVDTPGS--EDTMESRRSSMASTSTFF-----DTSSIGTVQ-----NPYA 459
Db 631 STATENPNSTEGSSANTNLEGSGQDADTGTGVVNNESQDTSDTGNAESGEQLQDSTQS 690
Qy 460 DVKTSLHDSQVPTNSNTSVQNMGTDSVVYSTIQHPPRDTT----DNGARLLGNPSAGI 515
Db 691 NEENTLPNSSIDQSNENTDESSDHSHEITDESVSSSKSGSSTPDQDGAASSGAPS-GD 749
Qy 516 QSTYAR--LALSGGLRHDMGGLTGGSNSAVNTSNPPAPGS 554
Db 750 QSISANACLAKSAASTDSSPVSNSGSDVTASSDNDPSSS 790

RESULT 7
US-09-620-412C-195
; Sequence 195, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
; APPLICANT: Steven P. Fling
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; FILE REFERENCE: 210121.469C7
; CURRENT APPLICATION NUMBER: US/09/620,412C
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 195
; LENGTH: 821
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-620-412C-195

Query Match
  5.8%; Score 163.5; DB 4; Length 821;
Best Local Similarity 21.9%; Pred. No. 0.00035;
Matches 127; Conservative 67; Mismatches 216; Indels 171; Gaps 27;

Qy 79 SEITLNDGFEVLHDHGLDPLTLNRQIGSSVFRVETQED--GK----- 117
```

```
Db 276 TEQTKSNG-----NQDGSSETKDTQVSESPESPDPDDVLGKGGIYTEKSLTITGITGTI 331
Qy 118 ----HIAVQRNGVETSVVLSDQEVARLQSDIDPEGKDFVFTGGRGGAGHAMVTVASDIT 173
Db 332 DFVSNIAATDSGAGVFTKENLSCTNTNSLQFLKN-----SAGQHGGG-AVVTQMSVT 382
Qy 174 EARQRIELLEPKGTGE-----SKGAGESKGVGELRESNGAENTTETOTSTSTSLR 226
Db 383 NTTSE--SITTPPLVGEVIFSENTAKHGCGGICTNKLSLN-----LKTVTLTKNSAKE 434
Qy 227 SDPKLWLALGTVATGLIGLAATGIVQALALTPEDSPPTTDP----- 268
Db 435 SG-----GAFTDLASIPITD-----TPESSTPSSSPASTPEVVASAKINRFPAS 480
Qy 269 ---DAAASATETATRDQL--TKEAFQNPDPNOKVNIDELGNAPSGVLKDDVVANIEEQA 323
Db 481 TAEPAAPSLTE-AESDQTDQETSDTNSDID-VSIENILN-----VAIQNTSAK 528
Qy 324 AA----GEEAKQQAENNAQAQKYDEQAQKQBELKVSSGAGYGLSGALLILG-----GG 374
Db 529 KGGAIYGGKAKLSRINN-----LELSGNSQDVGGGLCLITESVEFDA 570
Qy 375 IGVAVT-----AALHRKNQVPEQTTTTTTTTTTTSARTVEN---KPAANTPAQG-- 420
Db 571 IGSLLSHVNSAAKEGGVTHSKVTLSNLKSTFTPADNTVKAIVESTPEAPEIIPVEGEE 630
Qy 421 -----NVDTPGS--EDTMESRRSSMASTSTFF-----DTSSIGTVQ-----NPYA 459
Db 631 STATENPNSTEGSSANTNLEGSGQDADTGTGVVNNESQDTSDTGNAESGEQLQDSTQS 690
Qy 460 DVKTSLHDSQVPTNSNTSVQNMGTDSVVYSTIQHPPRDTT----DNGARLLGNPSAGI 515
Db 691 NEENTLPNSSIDQSNENTDESSDHSHEITDESVSSSKSGSSTPDQDGAASSGAPS-GD 749
Qy 516 QSTYAR--LALSGGLRHDMGGLTGGSNSAVNTSNPPAPGS 554
Db 750 QSISANACLAKSAASTDSSPVSNSGSDVTASSDNDPSSS 790

RESULT 8
US-09-598-419-195
; Sequence 195, Application US/09598419
; Patent No. 6565856
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
; FILE REFERENCE: 210121.469C6
; CURRENT APPLICATION NUMBER: US/09/598,419
; CURRENT FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 357
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 195
; LENGTH: 821
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-598-419-195

Query Match
  5.8%; Score 163.5; DB 4; Length 821;
Best Local Similarity 21.9%; Pred. No. 0.00035;
Matches 127; Conservative 67; Mismatches 216; Indels 171; Gaps 27;

Qy 79 SEITLNDGFEVLHDHGLDPLTLNRQIGSSVFRVETQED--GK----- 117
Db 276 TEQTKSNG-----NQDGSSETKDTQVSESPESPDPDDVLGKGGIYTEKSLTITGITGTI 331
Qy 118 ----HIAVQRNGVETSVVLSDQEVARLQSDIDPEGKDFVFTGGRGGAGHAMVTVASDIT 173
Db 332 DFVSNIAATDSGAGVFTKENLSCTNTNSLQFLKN-----SAGQHGGG-AVVTQMSVT 382
Qy 174 EARQRIELLEPKGTGE-----SKGAGESKGVGELRESNGAENTTETOTSTSTSLR 226
```



```

Db      383 NTTSE--SITTPPLVGEVIFSENTAKHGCGGICTNKLSLN-----LKTVTLTKNSAKE 434
QY      227 SPKMLWALGTWATGLIGLAAGTIVQALALTPPEPSPTTDP-----
Db      435 SG-----GAFTDLASIPITD-----TPESSTPSSSPASTPEVVASAKINRFFAS 480
QY      269 ---DAAASATETATRDQL--TKEAFONPDNOKVNDLGNALTPSGVLKDDVVANIEEOKAK 323
Db      481 TAEPAPASLITE-AESDQTDQETSDTNSDID-VSIENILN-----VAINQNTSAK 528
QY      324 AA-----GEEAKQAIENNAQAQKYDEOQAKRQEBELKVSSGAGYGLSGALILG-----GG 374
Db      529 KGGAIYKKAKLSRINN-----LELSGNSQDVGGGLCLTESVEFDA 570
QY      375 IGAVT-----AALHRKNQVPEQTITTTTTTTSARTVEN---KPANNTPAQG-- 420
Db      571 IGSLLSHYNSAAKEGGVIHKSIVTLNLKSTFTPADNTVKAIVESTPEAPEBIPPVEGEE 630
QY      421 -----NVDTPGS--EDTMESRRSSMASTSTFF-----DTSSIGTVQ-----NPYA 459
Db      631 STATENPNSNTEGSSANTNLESGQDGTADTGTGVVNNESQDTSDTGNAESGLODSTQS 690
QY      460 DVKTSLHDSQVPTSNSTSVQNMGNNTDSVYVSTIQHPPRDIT---DNGARLLGNPSAGI 515
Db      691 NEENTLPNSSIDQSNENTDESSDSTEETDESVSSSKSGSSTPDQCGAASSGAPS-GD 749
QY      516 QSTYAR--LALSGGLRHDMGGLTGGNSNAVNTSNPPAPGS 554
Db      750 QSISANACLAKEYAAASTDSSPVNSSGSDVTASSNDPSSS 790

```

RESULT 9

```

US-09-556-877-179
; Sequence 179, Application US/09556877
; Patent No. 6432916
; GENERAL INFORMATION:
; APPLICANT: Probet, Peter
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yaghi
; APPLICANT: Fling, Steve
; APPLICANT: Maisonneuve, Jeff
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C5
; CURRENT APPLICATION NUMBER: US/09/556,877
; CURRENT FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 179
; LENGTH: 1776
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-556-877-179

```

```

Query Match          5.8%; Score 163.5; DB+4; Length 1776;
Best Local Similarity 21.9%; Pred. No. 0.0011;
Matches 127; Conservative 67; Mismatches 216; Indels 171; Gaps 27;

QY      79 SEITLNDGFEVLHDGHPDITLNRQIGSSVFRVETOED--GK----- 117
Db      289 TEQTKSNG-----NQDGSSETKDTQVSESPESPDPDVLGKGGIYTEKSLTITGITGTI 344
QY      118 ---HIAVQQRNGVETSVVLSQDEYARLOSIDPEGKDFVFTGGRGGAGHAMVTVASDIT 173
Db      345 DFVSNIAITDSGAGVFTKENLSCNTNLSQFLKN-----SAGQHGGG-AVYVQTMVMT 395
QY      174 EARQRILELLEPKGTGE-----SKGAGESKGVGELRESNAGNTTETOTSTSTSLR 226
Db      396 NTTSE--SITTPPLVGEVIFSENTAKHGCGGICTNKLSLN-----LKTVTLTKNSAKE 447
QY      227 SDPKMLWALGTWATGLIGLAAGTIVQALALTPPEPSPTTDP----- 268

```

```

Db      448 SG-----GAFTDLASIPITD-----TPESSTPSSSPASTPEVVASAKINRFFAS 493
QY      269 ---DAAASATETATRDQL--TKEAFONPDNOKVNDLGNALTPSGVLKDDVVANIEEOKAK 323
Db      494 TAEPAPASLITE-AESDQTDQETSDTNSDID-VSIENILN-----VAINQNTSAK 541
QY      324 AA-----GEEAKQAIENNAQAQKYDEOQAKRQEBELKVSSGAGYGLSGALILG-----GG 374
Db      542 KGGAIYKKAKLSRINN-----LELSGNSQDVGGGLCLTESVEFDA 583
QY      375 IGAVT-----AALHRKNQVPEQTITTTTTTTSARTVEN---KPANNTPAQG-- 420
Db      584 IGSLLSHYNSAAKEGGVIHKSIVTLNLKSTFTPADNTVKAIVESTPEAPEBIPPVEGEE 643
QY      421 -----NVDTPGS--EDTMESRRSSMASTSTFF-----DTSSIGTVQ-----NPYA 459
Db      644 STATENPNSNTEGSSANTNLESGQDGTADTGTGVVNNESQDTSDTGNAESGLODSTQS 703
QY      460 DVKTSLHDSQVPTSNSTSVQNMGNNTDSVYVSTIQHPPRDIT---DNGARLLGNPSAGI 515
Db      704 NEENTLPNSSIDQSNENTDESSDSTEETDESVSSSKSGSSTPDQCGAASSGAPS-GD 762
QY      516 QSTYAR--LALSGGLRHDMGGLTGGNSNAVNTSNPPAPGS 554
Db      763 QSISANACLAKEYAAASTDSSPVNSSGSDVTASSNDPSSS 803

```

RESULT 10

```

US-09-620-412C-179
; Sequence 179, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
; APPLICANT: Steven P. Fling
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C7
; CURRENT APPLICATION NUMBER: US/09/620,412C
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 179
; LENGTH: 1776
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-620-412C-179

```

```

Query Match          5.8%; Score 163.5; DB 4; Length 1776;
Best Local Similarity 21.9%; Pred. No. 0.0011;
Matches 127; Conservative 67; Mismatches 216; Indels 171; Gaps 27;

QY      79 SEITLNDGFEVLHDGHPDITLNRQIGSSVFRVETOED--GK----- 117
Db      289 TEQTKSNG-----NQDGSSETKDTQVSESPESPDPDVLGKGGIYTEKSLTITGITGTI 344
QY      118 ---HIAVQQRNGVETSVVLSQDEYARLOSIDPEGKDFVFTGGRGGAGHAMVTVASDIT 173
Db      345 DFVSNIAITDSGAGVFTKENLSCNTNLSQFLKN-----SAGQHGGG-AVYVQTMVMT 395
QY      174 EARQRILELLEPKGTGE-----SKGAGESKGVGELRESNAGNTTETOTSTSTSLR 226
Db      396 NTTSE--SITTPPLVGEVIFSENTAKHGCGGICTNKLSLN-----LKTVTLTKNSAKE 447
QY      227 SDPKMLWALGTWATGLIGLAAGTIVQALALTPPEPSPTTDP----- 268
Db      448 SG-----GAFTDLASIPITD-----TPESSTPSSSPASTPEVVASAKINRFFAS 493
QY      269 ---DAAASATETATRDQL--TKEAFONPDNOKVNDLGNALTPSGVLKDDVVANIEEOKAK 323
Db      494 TAEPAPASLITE-AESDQTDQETSDTNSDID-VSIENILN-----VAINQNTSAK 541
QY      324 AA-----GEEAKQAIENNAQAQKYDEOQAKRQEBELKVSSGAGYGLSGALILG-----GG 374
Db      542 KGGAIYKKAKLSRINN-----LELSGNSQDVGGGLCLTESVEFDA 583

```

```
QY 375 IGAVT-----AALHRKNQPVQETTTTTTTTTTTTARTVEN---KPANNTPAQG-- 420
|| : : : : : : : : : : : : : : : : : : : : : : : :
Db 584 IGSLLSHYNSAAKEGGVIHSTKVTLSNLKSTFTFADNTVKAIVESTPEAPEIIPVEGEE 643
|| : : : : : : : : : : : : : : : : : : : : : : : :
QY 421 -----NVDTPGS--EDTMESRRSSMASTSTFF-----DTSSIGTVQ-----NPYA 459
|| : : : : : : : : : : : : : : : : : : : : : : : :
Db 644 STATENPNSNTGSSANTNLEGSGDGTADTGTGVVNNESQDTSDTGNAESGEQLQDSTQS 703
|| : : : : : : : : : : : : : : : : : : : : : : : :
QY 460 DVKTSLHDSQVPTNSNTSVQNMGNMNTDSVVYSTIQHPRPDRTT---DNGARLLGNPSSAGI 515
|| : : : : : : : : : : : : : : : : : : : : : : : :
Db 704 NEENTLPNSSIDQSNENTDESDSHTEITDESVSSSKSGSSTPDQDGAASSGAPS-GD 762
|| : : : : : : : : : : : : : : : : : : : : : : : :
QY 516 QSYAR--LALSGGLRHDMGLTGGNSAVNTSNPPAPGS 554
|| : : : : : : : : : : : : : : : : : : : : : : : :
Db 763 QSISANACLAASYAASTDSSPVNSSGSDVTASSDNPDS 803
|| : : : : : : : : : : : : : : : : : : : : : : : :
RESULT 11
US-09-598-419-179
; Sequence 179, Application US/09598419
; Patent No. 6565856
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Scholler, John
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C6
; CURRENT APPLICATION NUMBER: US/09/598,419
; CURRENT FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 357
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 179
; LENGTH: 1776
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-598-419-179
Query Match 5.8%; Score 163.5; DB 4; Length 1776;
Best Local Similarity 21.9%; Pred. No. 0.0011;
Matches 127; Conservative 67; Mismatches 216; Indels 171; Gaps 27;
QY 79 SEITLNDGFEVLHGHPLDNLNRQGTSSVFRVETOED--GK----- 117
|| : : : : : : : : : : : : : : : : : : : : : : : :
Db 289 TEQTKSNG---NQDGSSETKDQTVSESPESPDPDLVGKGGIYTEKSLTITGITGTI 344
|| : : : : : : : : : : : : : : : : : : : : : : : :
QY 118 ---HIAVQQRNGVETSVVLSDQEVARLQSIDPECKDFVFTGGRGGAGHAMVTVASDIT 173
|| : : : : : : : : : : : : : : : : : : : : : : : :
Db 345 DFVSNIAIDSGAGVFTKENLSCTNNSLQFLKN-----SAGQHGGG-AYVQTMTSVT 395
|| : : : : : : : : : : : : : : : : : : : : : : : :
QY 174 EARQRILELLEPKGTGE-----SKGAGESKGVGELRESNPGAENTTTETOTSTSTSLR 226
|| : : : : : : : : : : : : : : : : : : : : : : : :
Db 396 NITSE--SITTPPLVGEVIFSENTAKHGCGGICTNKLISLN-----LKTVTLTKNSAKE 447
|| : : : : : : : : : : : : : : : : : : : : : : : :
QY 227 SPKMLWALGTVATGLIGLAATGIVQALALALTEPPDSPPTTDP----- 268
|| : : : : : : : : : : : : : : : : : : : : : : : :
Db 448 SG-----GAFTDLASIPTD-----TPESSTPGSSSPASPEVVASAKINRPFAS 493
|| : : : : : : : : : : : : : : : : : : : : : : : :
QY 269 ---DAAASATEATRDQL--TKEAFQNPQNKVNTDELGNALPSGLVDVNVANIEEOAK 323
|| : : : : : : : : : : : : : : : : : : : : : : : :
Db 494 TAEPAAPEALTE-AESDQTDQETSDTNSDID-VSIENILN-----VAINQNTSAK 541
|| : : : : : : : : : : : : : : : : : : : : : : : :
QY 324 AA---GGEAKQQAENNAQAQKYDEQQAQKEELKVSSGAGYGLSGALLILG-----GG 374
|| : : : : : : : : : : : : : : : : : : : : : : : :
Db 542 KGAIIYGGKAKLSRNN-----LELSGNSSQVGGGLCTESVEFDA 583
|| : : : : : : : : : : : : : : : : : : : : : : : :
QY 375 IGAVT-----AALHRKNQPVQETTTTTTTTTTTTARTVEN---KPANNTPAQG-- 420
|| : : : : : : : : : : : : : : : : : : : : : : : :
Db 584 IGSLLSHYNSAAKEGGVIHSTKVTLSNLKSTFTFADNTVKAIVESTPEAPEIIPVEGEE 643
|| : : : : : : : : : : : : : : : : : : : : : : : :
QY 421 -----NVDTPGS--EDTMESRRSSMASTSTFF-----DTSSIGTVQ-----NPYA 459
|| : : : : : : : : : : : : : : : : : : : : : : : :
Db 644 STATENPNSNTGSSANTNLEGSGDGTADTGTGVVNNESQDTSDTGNAESGEQLQDSTQS 703
|| : : : : : : : : : : : : : : : : : : : : : : : :
```

```
QY 460 DVKTSLHDSQVPTNSNTSVQNMGNMNTDSVVYSTIQHPRPDRTT---DNGARLLGNPSSAGI 515
|| : : : : : : : : : : : : : : : : : : : : : : : :
Db 704 NEENTLPNSSIDQSNENTDESDSHTEITDESVSSSKSGSSTPDQDGAASSGAPS-GD 762
|| : : : : : : : : : : : : : : : : : : : : : : : :
QY 516 QSYAR--LALSGGLRHDMGLTGGNSAVNTSNPPAPGS 554
|| : : : : : : : : : : : : : : : : : : : : : : : :
Db 763 QSISANACLAASYAASTDSSPVNSSGSDVTASSDNPDS 803
|| : : : : : : : : : : : : : : : : : : : : : : : :
RESULT 12
US-09-248-796A-15936
; Sequence 15936, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 15936
; LENGTH: 441
; TYPE: PRT
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (322),(418)
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unknown
US-09-248-796A-15936
Query Match 5.6%; Score 159.5; DB 4; Length 441;
Best Local Similarity 21.9%; Pred. No. 0.00029;
Matches 90; Conservative 48; Mismatches 162; Indels 111; Gaps 15;
QY 183 LEPKGTGESKGAGSKGVGELRESNPGAENTTTETOTSTSTSLRSDPKLWALGTVATGL 242
|| : : : : : : : : : : : : : : : : : : : : : : : :
Db 88 LEKENTENNQTDNESAGVDEDNSTNKFDDIDAKSELSTSSSLRN-----GK 134
|| : : : : : : : : : : : : : : : : : : : : : : : :
QY 243 IGLAATGIVQALALTEPPDSPPTT-----TDPDAAASATETA---TRDQLTKEAFQNP 292
|| : : : : : : : : : : : : : : : : : : : : : : : :
Db 135 FSNLTLDLTKEMKKKSSPSLSNTENSLSVNNTNTASLSASNTLNKTPSKLDTASQSPN 194
|| : : : : : : : : : : : : : : : : : : : : : : : :
QY 293 NQKNIDELGNALPSGLVKDDVANI BEQAKAAGEEAKQQAENNAQAQKYDEQQAQK 352
|| : : : : : : : : : : : : : : : : : : : : : : : :
Db 195 TNSTLVSTFKKIGKWYFKDD-----GSSKKQHDEKLGYP 230
|| : : : : : : : : : : : : : : : : : : : : : : : :
QY 353 EELKVSSGAGVGLSGALLILGGIGVAVTAALHRKNQPVQETTTT-----TTTTTTSART 407
|| : : : : : : : : : : : : : : : : : : : : : : : :
Db 231 PEMTSSRRAPKSSSMTLGLSDSTRYTLATEENKDLPDIESANTTSASVNTATTAPSL 290
|| : : : : : : : : : : : : : : : : : : : : : : : :
QY 408 VENKANNTPAQGNVDTPGSEDPTMESRRSSMASTSTFFDTSSICTVQNPYADVK---TS 464
|| : : : : : : : : : : : : : : : : : : : : : : : :
Db 291 LSNK-GTTTAPVNVISLSNFTVPSVAVGADXTT-----TAPAPVANSKLTSTP 339
|| : : : : : : : : : : : : : : : : : : : : : : : :
QY 465 LHDSQVPTNSNTSVQNMGNMNTDSVVYSTIQH-----PPRDTTDN-----GAR 506
|| : : : : : : : : : : : : : : : : : : : : : : : :
Db 340 L-DNQTPSYNPSRLDNN--NEDLSIFSVAYQNAQAIIPSLTPPLPPRPDIISDLNLSIRSGS 396
|| : : : : : : : : : : : : : : : : : : : : : : : :
QY 507 LLGNPSAGIQSYARLALSGGLRHDMGGLTGGNSAVNT---SNNPPAPGS 554
|| : : : : : : : : : : : : : : : : : : : : : : : :
Db 397 LVGSEFTLSLQ-----PNSPIRKTKLXKPPPPPKS 426
|| : : : : : : : : : : : : : : : : : : : : : : : :
RESULT 13
US-09-248-796A-22338
; Sequence 22338, Application US/09248796A
; Patent No. 6747137
```

GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 22338
; LENGTH: 556
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-22338

Query Match 5.6%; Score 158; DB 4; Length 556;
Best Local Similarity 21.8%; Pred. No. 0.00054;
Matches 93; Conservative 44; Mismatches 200; Indels 90; Gaps 13;

QY 169 ASDITEARQRIELLEPKGTGSGKAGE-----SKGVGELRESNGAENT-----213
DB 71 STDITSASS--TDEQSSSGTGQSSTEDPIDSTESDTSSATSDSSATNTDNTNSES 128
QY 214 TETQSTSTSSLSRDPKLMALGATVATGLIGLAATGIVQA-----LALTP 258
DB 129 TDSSTATDSSSDTSSTASSTETNTDVTSDTSNTGATESSTATDNTDATDSSTVSETG 188
QY 259 EPDPTTTPDAAA-----SATETATRDQLTKEAFQNPNDKVNIDELGNAI 305
DB 189 ATDSSATDNTNGATSSSDTSNTGATESSTATDNTSATNTDNTGNTATNTD--NTAT 247
QY 306 PGVLKDDVVVIANIEQAKAGEAKQQAJENNAQAOKKYDEQAKRQBELKVSAGYGL 365
DB 248 DTSSSTETATNTD-----GTETNGATETNTDTSASNTDNDT-----GS 287
QY 366 SGALLIGGGIGVAVTAALHRKQNPVQVTTTTTTTTTTSARTVENKPAANTPAQGNVDP 425
DB 288 NTATNTGG-----TDNTDNTGTDNTGTDNTGTDNTGTDNTGTDNTGTDNTG 329
QY 426 GSEDVTMESRRSSMASTSTFFDTSSIGTVQNPY--ADVKTSLHDSQVPTSNSTSVQNG 483
DB 330 NMGATETATNTNGNTNGTNTGATDNTATNTGATDNTGATDNTGATDNTGATDNTGAT 389
QY 484 NTDSVVYSTIQHPPRDT--TDNGARLLGNPSAGIQSTYARLALSG--GLRHDMGGLTGGSNS 541
DB 390 NTDNV--STTKDIPSPSTINEGSGNGSGNGSGNGSGNGSGNGSGNGSGNGSGNGSGNG 447
QY 542 AVNTSN 548
DB 448 SCNGSDN 454

RESULT 14
US-09-248-796A-17307
; Sequence 17307, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 17307
; LENGTH: 827

TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-17307

Query Match 5.5%; Score 157.5; DB 4; Length 827;
Best Local Similarity 22.5%; Pred. No. 0.0011;
Matches 126; Conservative 63; Mismatches 251; Indels 119; Gaps 24;

QY 56 ADSGNRASDVGLPVNPMRLAASEITLNDGFEVLH--DHGPLDTLNRQIGSVFVETQ 113
DB 37 ADNIDKRA---GAIGNFFRDFTNIFGNDNLEVNQPNSTNGATST--GHFFGSPSTSTH 91
QY 114 EDGKHIAVGORNGVTSVVLSDQEVARLQSIDPEGKQKVFETGGRGAGHAMVTVASDIT 173
DB 92 -----QQTPGTISNNVNTKSSQNOQS--PS-----TSPSTVAAAATSSPVA 133
QY 174 EARQRIELLEPKGTGSGKAGESGVGELRESNGAENTTTTQSTSTSSLSRSDPKML 233
DB 134 STR-----PASTSEQKQEESTA---RQSTSPATTTTNTPTTSSPSTSKETPTSN 181
QY 234 ALGTVA-----TGLIGLAATGIV-----QALALTP-----EPDPTTTPDAAA 272
DB 182 AQTSSANNQOQSNTAAPSTSVIQPSTSEVHVQSQQTSTTPTNTPTSSPNTPTTSEAPT 241
QY 273 SATETATRDQLTKEAFQNPNDKV-----NIDELGNA-----IPSGVLKDDVVANIEEQ 322
DB 242 SAAPT-----TSEAPVTPSTSEVNPNTTSEAPNTPTTSEAPVTPSTSEVNPNTPTS 295
QY 323 KAAGEAKQKQ--AIENNAQAOKKYDEQQA---KQBELKVSAGYGLSGALILGGGIG 376
DB 296 KAPNTPTTSEAPATTTTSEAPNTPTTSEAPVTPPTTSEVNPNTTSTQGDVSTS-----S 348
QY 377 VAVTAALHRKQNPVQVTTTTTTTTTTSARTVENKPA---NTPA-----QGNVD 423
DB 349 TSVTEQTTLTSTLLPPTTALTQTSTPEASDPKPSSTSIETPSTSTFEQDPTTTSSVG 408
QY 424 TPGSEDTMESRRSSMASTSTFFDTSSIGTVQNPYADVKTSLHDSQVPTSNSTSVQNG 483
DB 409 TPSSPOPTTTSELAVTSNS--PTQESTSLVEP---TTSSLESNTPTNPSTSEAPQS 463
QY 484 NTDSVVYSTIQHPPRDTTDNGARLLGNPSAGIQSTYARLALSGGL-----RHDMGGLTG 537
DB 464 TSAS-----QAPPTTSSAPAPELSSNADPSNLVLSHSETTSLVNPSTDQIDSSSTD 517
QY 538 GNSAVN--TSNNPPAGS 554
DB 518 AVSQATTEPTSENTPTAAS 536

RESULT 15
US-09-949-016-10491
; Sequence 10491, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10491
; LENGTH: 2045
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10491

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 13, 2005, 11:16:56 ; Search time 178 Seconds
(without alignments)
1605.282 Million cell updates/sec

Title: US-09-189-415b-11

Perfect score: 2840

Sequence: 1 MPIGNLGHNPVNSIPPAP.....SNSAVNTSNPPAPGSHRFV 558

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2840	100.0	558	2	Q7DB77	Q7db77 escherichia
2	2840	100.0	558	2	Q9R396	Q9r396 escherichia
3	2824	99.4	558	2	O85506	O85506 escherichia
4	1843	64.9	538	2	Q47014	Q47014 escherichia
5	1835	64.6	538	2	Q47016	Q47016 escherichia
6	1824	64.2	538	2	O85508	O85508 escherichia
7	1613.5	56.8	551	2	O68258	O68258 escherichia
8	1594	56.1	552	2	Q9KWH9	Q9kwh9 escherichia
9	1585	55.8	550	2	O52147	O52147 escherichia
10	1569.5	55.3	547	2	Q7BHL5	Q7bhl5 escherichia
11	1569.5	55.3	547	2	Q9WXK1	Q9wxk1 escherichia
12	1568.5	55.2	547	2	Q9ETI1	Q9eti1 citrobacter
13	1531.5	53.9	549	2	O50190	O50190 escherichia
14	1146.5	40.4	367	2	Q79C12	Q79c12 escherichia
15	215.5	7.6	1374	2	Q7YU77	Q7yu77 drosophila
16	214.5	7.6	1323	2	Q7KSZ0	Q7kszo drosophila
17	214.5	7.6	1376	2	Q9VI63	Q9vi63 drosophila
18	208.5	7.3	1321	2	Q962D1	Q962d1 drosophila
19	207.5	7.3	1015	2	Q8SZW9	Q8szw9 drosophila
20	200	7.0	2332	2	Q81FX6	Q81fx6 caenorhabdi
21	196	6.9	2338	2	Q759S3	Q759s3 ashbya goss
22	195.5	6.9	1254	2	Q94185	Q94185 caenorhabdi
23	194.5	6.8	3295	2	O66GT3	O66gt3 rattus norv
24	187.5	6.6	3550	2	O66GT4	O66gt4 rattus norv
25	187	6.6	1291	2	Q77261	Q77261 drosophila
26	187	6.6	1300	2	Q9W5E0	Q9w5e0 drosophila
27	185	6.5	2310	2	Q8CMU7	Q8cmu7 staphylococ
28	184.5	6.5	3135	2	Q7KHU2	Q7khu2 drosophila
29	184	6.5	1961	2	Q7PQ78	Q7pq78 anopheles g
30	182	6.4	574	2	Q7R861	Q7r861 neurospora
31	182	6.4	3080	2	Q9V602	Q9v602 drosophila

32	182	6.4	3109	2	Q9BMO0	Q9bmq0 drosophila
33	180	6.3	596	2	Q6UXC5	Q6uxc5 homo sapien
34	178.5	6.3	3443	2	Q6JZM8	Q6jzm8 mus musculu
35	177	6.2	1351	2	Q6GDE9	Q6gde9 staphylococ
36	176	6.2	2283	2	Q8VQ99	Q8vg99 staphylococ
37	174.5	6.1	1241	2	Q75JCO	Q75jco dictyosteli
38	174.5	6.1	2634	2	O63KC6	O63kc6 burkholderi
39	174	6.1	1176	2	Q869T7	Q869t7 dictyosteli
40	174	6.1	1535	2	Q62D27	Q62d27 burkholderi
41	174	6.1	2275	2	Q8NUJ3	Q8nuj3 staphylococ
42	174	6.1	2275	2	O6G620	O6g620 staphylococ
43	174	6.1	5899	2	Q74GG7	Q74gg7 geobacter s
44	173	6.1	1376	2	Q7SSH8	Q7ssh8 neurospora
45	172.5	6.1	2271	2	Q99QY4	Q99qy4 staphylococ

ALIGNMENTS

RESULT 1

Q7DB77	PRELIMINARY;	PRT;	558 AA.
AC	Q7DB77; Q7A9Q1;		
DT	05-JUL-2004 (Tremblrel. 27, Created)		
DT	05-JUL-2004 (Tremblrel. 27, Last sequence update)		
DT	25-OCT-2004 (Tremblrel. 28, Last annotation update)		
DE	Putative translocated intimin receptor protein (Translocated intimin receptor Tir).		
GN	Name=tir; OrderedLocusNames=ECs4561, Z5112;		
OS	Escherichia coli O157:H7.		
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;		
OC	Enterobacteriaceae; Escherichia.		
OX	NCBI_TaxID=83334;		
RN	[1]		
RC	SEQUENCE FROM N.A.		
RC	STRAIN=O157:H7 / EDL93 / ATCC 700927 / EHEC;		
RX	MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;		
RA	Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,		
RA	Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,		
RA	Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,		
RA	Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamoukis K.,		
RA	Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,		
RA	Welch R.A., Blattner F.R.;		
RT	"Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.;"		
RL	Nature 409:529-533(2001).		
RN	[2]		
RC	SEQUENCE FROM N.A.		
RC	STRAIN=O157:H7 / RIMD 050952 / EHEC;		
RX	MEDLINE=21156231; PubMed=11258796;		
RA	Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,		
RA	Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,		
RA	Iida T., Takami H., Honda T., Sasaki C., Ogasawara N., Yasunaga T.,		
RA	Kuhara S., Shiba T., Hattori M., Shinagawa H.;		
RT	"Complete genome sequence of enterohemorrhagic Escherichia coli		
RT	O157:H7 and genomic comparison with a laboratory strain K-12.;"		
RL	DNA Res. 8:11-22(2001).		
DR	EMBL; AE005595; AAG58825.1; --		
DR	EMBL; AP002566; BAB37984.1; --		
DR	GO; GO:0004872; F:receptor activity; IEA.		
DR	GO; GO:0007155; P:cell adhesion; IEA.		
DR	InterPro; IPR003536; Tir_receptor.		
DR	Pfam; PF07489; Tir_receptor C; 1.		
DR	Pfam; PF03549; Tir_receptor M; 1.		
DR	Pfam; PF07490; Tir_receptor N; 1.		
DR	PRINTS; PR01370; TRANSINTINR.		
SQ	SEQUENCE 558 AA; 58022 MW; 99C41722D4B4AA1 CRC64;		

Query Match 100.0%; Score 2840; DB 2; Length 558;

Best Local Similarity 100.0%; Pred. No. 3e-142;

Matches 558; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPIGNLGHNPVNSIPPAPLPSCDTGAGRGQLINSTGPIGSRALTPVNSVADSGD 60

```
Db 1 MPIGNLGNPNVNNISIPAPLPSTQDAGGARGQLINSTPLGSRALFTPVVNSMADSGD 60
QY 61 NRASDVPLGPNVPMRLAAASEITLNDGFEVLHDHGPLDTLNROI GSSVFRVETQEDGKHIA 120
Db 61 NRASDVPLGPNVPMRLAAASEITLNDGFEVLHDHGPLDTLNROI GSSVFRVETQEDGKHIA 120
QY 121 VQORNGVETSVVLSDOEYARLQSIDPEGKDKFVFTGGRGGAGHAMVTVASDITEARQRI 180
Db 121 VQORNGVETSVVLSDOEYARLQSIDPEGKDKFVFTGGRGGAGHAMVTVASDITEARQRI 180
QY 181 ELLEPKGTGESKGAGESKVGELRESNLSGAENTTETQTSSTSLRSDPKLWALGTAT 240
Db 181 ELLEPKGTGESKGAGESKVGELRESNLSGAENTTETQTSSTSLRSDPKLWALGTAT 240
QY 241 GLIGLAATGIVQALALTPEDPSPTTDPDAAASATETATRDOLTKEAFQNDKQVNI 300
Db 241 GLIGLAATGIVQALALTPEDPSPTTDPDAAASATETATRDOLTKEAFQNDKQVNI 300
QY 301 LGNAIPSGVLKDDVVANIEEQAAGEAKQAQAIENNAQAQKYDEQAQKEELKVSSG 360
Db 301 LGNAIPSGVLKDDVVANIEEQAAGEAKQAQAIENNAQAQKYDEQAQKEELKVSSG 360
QY 361 AGYGLSGALILGGGIGVAVTAALHRKNQPVETTTTTTTTTTSARTVENKPANNTPAQ 420
Db 361 AGYGLSGALILGGGIGVAVTAALHRKNQPVETTTTTTTTTTSARTVENKPANNTPAQ 420
QY 421 NVDTFGSEDTMESRRSSMASTSTFTDSSIGTVQNPYADVKTSLHDSQVPTSNTSVQ 480
Db 421 NVDTFGSEDTMESRRSSMASTSTFTDSSIGTVQNPYADVKTSLHDSQVPTSNTSVQ 480
QY 481 NMGNTDSVVYSTIOHPPRDITDNGARLLGNPSAGIQSTYARLALSGGLRHDWGLTGGSN 540
Db 481 NMGNTDSVVYSTIOHPPRDITDNGARLLGNPSAGIQSTYARLALSGGLRHDWGLTGGSN 540
QY 541 SAVNTSNPPAPGSHRFV 558
Db 541 SAVNTSNPPAPGSHRFV 558

RESULT 2
Q9R396 PRELIMINARY; PRT; 558 AA.
AC Q9R396;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Translocated intimin receptor Tir (L0027).
GN Name=tir;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=86/24;
RX MEDLINE=99242825; PubMed=10225900;
RA Devlinney B., Stein M., Reinscheid D., Abe A., Ruschkowski S.,
RA Finlay B.B.; Blattner F.R.;
RT "Enterohemorrhagic Escherichia coli O157:H7 produces Tir, which is
RT translocated to the host cell membrane but is not tyrosine
RT phosphorylated.";
RL Infect. Immun. 67:2389-2398(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC43895, and EDL933;
RX MEDLINE=98339885; PubMed=9673266;
RA Perna N.T., Mayhew G.F., Posfai G., Elliott S., Donnenberg M.S.,
RA Kaper J.B., Blattner F.R.;
RT "Molecular evolution of a pathogenicity island from enterohemorrhagic
RT Escherichia coli O157:H7";
RL Infect. Immun. 66:3810-3817(1998).
DR EMBL; AF125993; AAD29391.1; -.
DR EMBL; AF071034; AAC31506.1; -.

```

```
PIR; A98199; A98199.
PIR; E86045; E86045.
DR HSSP; Q9KWH9; 1P02.
DR GO; GO:0004872; P:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003536; Tir_receptor.
DR Pfam; PF07489; Tir_receptor C; 1.
DR Pfam; PF07490; Tir_receptor N; 1.
DR PRINTS; PR01370; TRNSINTIMR.
KW Receptor.
SQ SEQUENCE 558 AA; 58022 MW; 99C417222D4B4AA1 CRC64;

Query Match 100.0%; Score 2840; DB 2; Length 558;
Best Local Similarity 100.0%; Pred. No. 3e-142;
Matches 558; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPIGNLGNPNVNNISIPAPLPSTQDAGGARGQLINSTPLGSRALFTPVVNSMADSGD 60
Db 1 MPIGNLGNPNVNNISIPAPLPSTQDAGGARGQLINSTPLGSRALFTPVVNSMADSGD 60
QY 61 NRASDVPLGPNVPMRLAAASEITLNDGFEVLHDHGPLDTLNROI GSSVFRVETQEDGKHIA 120
Db 61 NRASDVPLGPNVPMRLAAASEITLNDGFEVLHDHGPLDTLNROI GSSVFRVETQEDGKHIA 120
QY 121 VQORNGVETSVVLSDOEYARLQSIDPEGKDKFVFTGGRGGAGHAMVTVASDITEARQRI 180
Db 121 VQORNGVETSVVLSDOEYARLQSIDPEGKDKFVFTGGRGGAGHAMVTVASDITEARQRI 180
QY 181 ELLEPKGTGESKGAGESKVGELRESNLSGAENTTETQTSSTSLRSDPKLWALGTAT 240
Db 181 ELLEPKGTGESKGAGESKVGELRESNLSGAENTTETQTSSTSLRSDPKLWALGTAT 240
QY 241 GLIGLAATGIVQALALTPEDPSPTTDPDAAASATETATRDOLTKEAFQNDKQVNI 300
Db 241 GLIGLAATGIVQALALTPEDPSPTTDPDAAASATETATRDOLTKEAFQNDKQVNI 300
QY 301 LGNAIPSGVLKDDVVANIEEQAAGEAKQAQAIENNAQAQKYDEQAQKEELKVSSG 360
Db 301 LGNAIPSGVLKDDVVANIEEQAAGEAKQAQAIENNAQAQKYDEQAQKEELKVSSG 360
QY 361 AGYGLSGALILGGGIGVAVTAALHRKNQPVETTTTTTTTTTSARTVENKPANNTPAQ 420
Db 361 AGYGLSGALILGGGIGVAVTAALHRKNQPVETTTTTTTTTTSARTVENKPANNTPAQ 420
QY 421 NVDTFGSEDTMESRRSSMASTSTFTDSSIGTVQNPYADVKTSLHDSQVPTSNTSVQ 480
Db 421 NVDTFGSEDTMESRRSSMASTSTFTDSSIGTVQNPYADVKTSLHDSQVPTSNTSVQ 480
QY 481 NMGNTDSVVYSTIOHPPRDITDNGARLLGNPSAGIQSTYARLALSGGLRHDWGLTGGSN 540
Db 481 NMGNTDSVVYSTIOHPPRDITDNGARLLGNPSAGIQSTYARLALSGGLRHDWGLTGGSN 540
QY 541 SAVNTSNPPAPGSHRFV 558
Db 541 SAVNTSNPPAPGSHRFV 558

RESULT 3
O85506 PRELIMINARY; PRT; 558 AA.
AC O85506;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DE Translocated intimin receptor Tir.
GN Name=tir;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=95SF2;
RX MEDLINE=99003184; PubMed=9784578;
RA Paton A.W., Manning P.A., Woodrow M.C., Paton J.C.;
RT "Translocated intimin receptors (tir) of Shiga-toxicogenic Escherichia
RT coli isolates belonging to serogroups O26, O111, and O157 react with
RT sera from patients with hemolytic-uremic syndrome and exhibit marked
RT sequence heterogeneity.";
RL Infect. Immun. 66:5580-5586 (1998).
DR EMBL; AF070067; AAC69314.1; -.
DR HSSP; Q9KWH9; 1F02.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003536; Tir_receptor.
DR Pfam; PF07489; Tir_receptor_C; 1.
DR Pfam; PF03549; Tir_receptor_M; 1.
DR Pfam; PF07490; Tir_receptor_N; 1.
DR PRINTS; PR01370; TRNSINTIMINR.
KW Receptor.
SQ SEQUENCE 558 AA; 58175 MW; CA2CDDAC94527C2E CRC64;

Query Match 99.4%; Score 2824; DB 2; Length 558;
Best Local Similarity 99.5%; Pred. No. 2.1e-141;
Matches 555; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MPIGNLGNPNVNNISIPAPPLPSQTDGAGRGQLINSTGLGSRALFTPVNSMADSG 60
DB 1 MPIGNLGNPNVNNISIPAPPLPSQTDGAGRGQLINSTGLGSRALFTPVNSMADSG 60
QY 61 NRASDVPLGLPVNPMRLAASEITLNDGFVLHDHGPLDTLNRQIGSSVFRVTOEDGKHIA 120
DB 61 NRASDVPLGLPVNPMRLAASEITLNDGFVLHDHGPLDTLNRQIGSSVFRVTOEDGKHIA 120
QY 121 VQORNGVETSVLSDQYARLOSIDPEGKDKVFTGGRGAGHAMVTVASDITEARQRL 180
DB 121 VQORNGVETSVLSDQYARLOSIDPEGKDKVFTGGRGAGHAMVTVASDITEARQRL 180
QY 181 ELLEPKGTGSKGAGSKGVGELRENSGAENTTETQSTSSLSRSDPKLWALGTAT 240
DB 181 ELLEPKGTGSKGAGSKGVGELRENSGAENTTETQSTSSLSRSDPKLWALGTAT 240
QY 241 GLIGLAATGIVQALALTPEDSPPTTDPDAASATETATRDQLTKFAQNPQKNVIDE 300
DB 241 GLIGLAATGIVQALALTPEDSPPTTDPDAASATETATRDQLTKFAQNPQKNVIDE 300
QY 301 LGNAIPSGVLKDDVANIIEQAKAAGEAKQAENNAQAQKYDEQQAQKQBELKVSSG 360
DB 301 LGNAIPSGVLKDDVANIIEQAKAAGEAKQAENNAQAQKYDEQQAQKQBELKVSSG 360
QY 361 AGVGLSGALILGGIGVAVTAALHRKNQVPEQTITTTTTTTTTTSARTVENKPNNTPAQ 420
DB 361 AGVGLSGALILGGIGVAVTAALHRKNQVPEQTITTTTTTTTTTSARTVENKPNNTPAQ 420
QY 421 NVDTPGSEDTMESRRSSMASTSTFPDTSIGTVQNPYADVKTSLHDSQVPTSNSNTSVQ 480
DB 421 NVDTPGSEDTMESRRSSMASTSTFPDTSIGTVQNPYADVKTSLHDSQVPTSNSNTSVQ 480
QY 481 NMGTDSVVYTIQHPRDTTNGARLLGNPAGIOSTYARLALSGGLHDMGGLTGGSN 540
DB 481 NMGTDSVVYTIQHPRDTTNGARLLGNPAGIOSTYARLALSGGLHDMGGLTGGSN 540
QY 541 SAVNTSNPPAGSHRFV 558
DB 541 SAVNTSNPPAGSHRFV 558

RESULT 4
ID Q47014 PRELIMINARY; PRT; 538 AA.
AC Q47014;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Translocated intimin receptor tir (Translocated intimin co-receptor)

```

```

DE GN (EspE protein).
OS Name=tir; Synonyms=espE;
OC Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1] SEQUENCE FROM N.A.
RC STRAIN=REPEC 84/110/1, and E65/56;
RA Krejany E.O.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RC SEQUENCE FROM N.A.
RX STRAIN=95ZG1;
RY MEDLINE=99003184; PubMed=9784578;
RA Paton A.W., Manning P.A., Woodrow M.C., Paton J.C.;
RT "Translocated intimin receptors (tir) of Shiga-toxicogenic Escherichia
RT coli isolates belonging to serogroups O26, O111, and O157 react with
RT sera from patients with hemolytic-uremic syndrome and exhibit marked
RT sequence heterogeneity.";
RL Infect. Immun. 66:5580-5586 (1998).
RN [3]
RC SEQUENCE FROM N.A.
RX STRAIN=B10;
RY MEDLINE=20187493; PubMed=10722617;
RX DOI=10.1128/IAI.68.4.2171-2182.2000;
RA Marches O., Nougayrede J.P., Boullier S., Mainil J., Charlier G.,
RA Raymond I., Pohl P., Boury M., De Rycke J., Milton A., Oswald E.;
RT "Role of tir and intimin in the virulence of rabbit enteropathogenic
RT Escherichia coli serotype O103:H2.";
RL Infect. Immun. 68:2171-2182 (2000).
RN [4]
RC SEQUENCE FROM N.A.
RX STRAIN=413/89-1;
RY MEDLINE=982994040; PubMed=9632251;
RA Deibel C., Kraemer S., Chakraborty T., Ebel F.;
RT "EspE, a novel secreted protein of attaching and effacing bacteria, is
RT tyrosine-phosphorylated 90 kDa protein.";
RL Mol. Microbiol. 28:463-474 (1998).
RN [5]
RC SEQUENCE FROM N.A.
RX STRAIN=413/89-1;
RY Benkel P., Chakraborty T.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; U59502; AAC32028.2; -.
DR EMBL; AF070068; AAC69316.1; -.
DR EMBL; AF113597; AAF03080.1; -.
DR EMBL; AF132728; AAD27868.1; -.
DR EMBL; AJ223063; CAA11065.1; -.
DR EMBL; AJ277443; CAC81869.1; -.
DR HSSP; Q9KWH9; 1F02.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003536; Tir_receptor.
DR Pfam; PF07489; Tir_receptor_C; 1.
DR Pfam; PF03549; Tir_receptor_M; 1.
DR Pfam; PF07490; Tir_receptor_N; 1.
DR PRINTS; PR01370; TRNSINTIMINR.
KW Receptor.
SQ SEQUENCE 538 AA; 55420 MW; 31D7A8E227B3D06C CRC64;

Query Match 64.9%; Score 1843; DB 2; Length 538;
Best Local Similarity 65.9%; Pred. No. 1.5e-89;
Matches 371; Conservative 61; Mismatches 101; Indels 30; Gaps 6;

```

```

QY 1 MPIGNLGNPNVNNISIPAPPLPSQTDGAGRGQLINSTGLGSRALFTPVNSMADSG 59
DB 1 MPIGNLGNPNVNNISIPAPPLPSQTDGAGRGQLINSTGLGSRALFTPVNSMADSG 60
QY 60 NRASDVPLGLPVNPMRLAASEITLNDGFVLHDHGPLDTLNRQIGSSVFRVTOEDGKHIA 119
DB 61 DSRASDIIFGLPTNPLRFRASEVSLHGALEVLHDKGLDGLTNSAIGSSLPFRVTRDGGSHV 120

```


Qy	420	GNVDTPGSEDTWESRRSSMASTSTFFDTSSIGTVQNPYADVKTSLHDSQVPTSNSTSV	479
Db	404	NTTDTSGPEESPASRRNSASLNAGSDTSTGTGTENPYADV-----GMPRNDSLARI	456
Qy	480	QNMGNMTDSVW---YSTIQHPRDPTDNGARLLGNPNSAGIQSTYARLALSGGLRHDGMGL	535
Db	457	SEEPYDEVAADPNYSVLQHSNGNSPVTG-RLVGTGPGQIQSTYALLASSGGLRLGMGL	515
Qy	536	TGGSNSAVNTSNPPAPGSHRPV 558	
Db	516	TGGCESAVSTANAAPTGPGRFV 538	

RESULT 6

OB5508 PRELIMINARY; PRT; 538 AA.

ID OB5508

AC OB5508;

DT 01-NOV-1998 (TReMBLrel. 08, Created)

DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)

DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)

DE Translocated intimin receptor Tir.

GN Name=tir;

OS Escherichia coli.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Escherichia.

NCBI_TaxID=562;

RN [1]

RC SEQUENCE FROM N.A.

RP STRAIN=EPEC87A;

RC MEDLINE=99003184; PubMed=9784578;

RX Paton A.W., Manning P.A., Woodrow M.C., Paton J.C.;

RT "Translocated intimin receptors (Tir) of Shiga-toxicogenic Escherichia

RT coli isolates belonging to serogroups O26, O111, and O157 react with

RT sera from patients with hemolytic-uremic syndrome and exhibit marked

RT sequence heterogeneity.";

RL Infect. Immun. 66:5580-5586(1998).

DR EMBL; AF070069; AAC69318.1; -.

DR HSPSP; Q9KWH9; 1F02.

DR GO; GO:0004872; P:receptor activity; IEA.

DR GO; GO:0007155; P:cell adhesion; IEA.

DR InterPro: IPR003536; Tir receptor.

DR Pfam; PF07489; tir_receptor_C; 1.

DR Pfam; PF03549; Tir_receptor_M; 1.

DR Pfam; PF07490; Tir_receptor_N; 1.

DR PRINTS; PR01370; TRNSINTMINR.

KW Receptor.

SQ SEQUENCE 538 AA; 55602 MW; 447052A0E3214D6D CRC64;

Query Match 64.2%; Score 1824; DB 2; Length 538;

Best Local Similarity 65.8%; Pred. No. 1.5e-88;

Matches 369; Conservative 60; Mismatches 106; Indels 26; Gaps 6

Qy	1	MPIGNLGNHPNVNNSIPAPPLPSQTDGAGG-RGQILNSTGSLGRALFTPVNSMADSG	59
Db	1	MPIGNLGNHPNVRALIPAPPLPSQTDGAGGARNQLINSGPMGSRLLFTPIRNSVADAA	60
Qy	60	DNRASDVPLGPNVPMRLAASETTLLNDGPEVLVHDHGFLDTLNRQIGSSVFRVETQSDGKH	119
Db	61	DSRASDIPLGTPNPLRFAASEVLSHGALEVLVHDKGLDTLNSAIGSSLFRVETRDGSHV	120
Qy	120	AVGQRNGVETSVLSDQRYARLQSDIPDGKOKFVFTGGRGGAGHAMVTVASDITEARQRI	179
Db	121	AIGQKNGLETTVVLSDQEFFSISQSLDPEGKKNFVFTGGRGGAGHAMVTVASDIAEARQRI	180
Qy	180	LELLEPKGTGSGKAGESKGVGELRESNSGAENTTTTQSTSTSSLRSDPKLWALGTVA	239
Db	181	IDKLEPKDTKETKEFG-----DPNSGEGKIIETHTSTFTSLRADPKLWLSLGTIA	231
Qy	240	TGLIGLAATGTVQALALTPEDSPPTTDPDAAASATATATRDOLTKFAFONPDNOKVND	299
Db	232	AGLIGMAATGTAQVALVTEPDDPTTTPDPTAASTAEATKDRLLTQEAQDPDPKQKVID	291
Qy	300	ELGNAIPSGVLKDDVVANIEEQAKAAGEAQQAENNAQAOKKYDEQAQKEBELKVSS	359


```

Db 1 MPIGNLGNVNGNHLIPAPLPSQTDGAARGGTGHLISSTGALGSRSLFSLRNSMADS 60
QY 59 GNRASDVPLGPNMRLAA--SEITLNDGFEVLHDHGPDLTLNRQIGSSVFRVETQBDG 116
Db 61 VDSR--DIPGLPTNPSRLAAATSETCLLGGFEVLHDHGPDLTLNRQIGSPAFRVEVQADG 118
QY 117 KHIAVGORNGVTSVVLSDQEVARLQSIDPEGKDKFVFTGGRGGAGHAMVTVASDITEAR 176
Db 119 THAAIGKNGLEVSVTLSPQEWSSLSQSIDTEGKRFVFTGGRGGGHPMTVTVASDIAEAR 178
QY 177 QRILELLEPKGTG-----ESKGAGESKGVGELRESNGAENTTETOTSTSTSSLRSDP 227
Db 179 TKILAKLDPDNHGGRQPKDVTDRSVGVSASGI-----DDGV--VSETHTSTTSSVRS 230
QY 228 DKLWALGTATGLIGLAATGIVQALALTPEDSPSTTTDPDAASATATATRDQLTKEA 287
Db 231 DPKFWVSGAIAAGLAGLAATGIAQALALTPEDDPPTTTDPDAANAASATKDQLTQEA 290
QY 288 FQNPONKVNIDELGNAPSGVLKDDVVANIEEOAKAAGEAKQQAIAENNAOAKKYDEQ 347
Db 291 FQNPENKVNIDANGNAIPSGELKDDIVEQIAQQAKEAGEVARQQAQVESNAQAQRYEDQ 350
QY 348 QAKRQBELKVSSGAGYGLSGALILGGGIGVAVTAALHRKNQPVQETTTTTTTTTSART 407
Db 351 HARRQBELQLSSGIGYGLSSALIVAGGIGAGVTTALHRRNQPAEQTTTTTHT-----V 404
QY 408 VENKPAANTPAQGNVDTPGSEDTMESRRSSMASTSTSTPTDTSIGTVQNPYADV-----K 462
Db 405 VQOQTGGNTPAQGGTDAETRAEDASLNRDSQGSVASTHWSDS--SEVNPYAEVGGARNS 463
QY 463 TSLHD-----SQVPTSNSTSVQNMGTDSVVYSTIOHPRPDITDNGARLGNPSAGIQ 516
Db 464 LSAHQPEEHIYDEVAADPGYSVIONFSGGPVT-----GRLIGTPQGQIGQ 508
QY 517 STYARLALSGGLRHDHMGGLTGGSNSAVNTSNPPAPGSHRFV 558
Db 509 STYALLANSGGLRLGMLGGLTGGSNSAVNSNAATPGVRFV 550

RESULT 10
Q7BHL5
ID AC Q7BHL5 PRELIMINARY; PRT; 547 AA.
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
DE Translocated intimin receptor Tir.
OS Citrobacter rodentium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Citrobacter.
OX NCBI_TaxID=67825;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DS100;
RX MEDLINE=21437640; PubMed=11553577;
RX DOI=10.1128/IAI.69.10.6323-6335.2001;
RA Deng W., Li Y., Vallance B.A., Finlay B.B.;
RT "Locus of enterocyte effacement from Citrobacter rodentium: sequence
RT analysis and evidence for horizontal transfer among attaching and
RT effacing pathogens.";
RL Infect Immun. 69:6323-6335(2001).
DR EMBL; AF311901; AAL08376.1; -.
DR GO; GO:0004872; Fireceptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003536; Tir_receptor.
DR Pfam; PF07489; Tir_receptor C; 1.
DR Pfam; PF03549; Tir_receptor M; 1.
DR Pfam; PF07490; Tir_receptor N; 1.
DR PRINTS; PR01370; TRNSINTMINR.
KW Receptor.
SQ SEQUENCE 547 AA; 56283 MW; 02CAC6D625FA6EE1 CRC64;

```

Query Match 55.3%; Score 1569.5; DB 2; Length 547;

```

Best Local Similarity 57.0%; Pred. No. 4.5e-75;
Matches 332; Conservative 69; Mismatches 122; Indels 59; Gaps 11;

QY 1 MPIGNLGNVNNNSIPPAPLPSQTDGA--GGRQQLINSTGLGSRALFTFVRNSMADS 58
Db 1 MPIGNLGNNNISNNLIIPAPLPSQTDGATRGNGSSLSSTGSLGSRLLFSLRSSIVDT 60
QY 59 GNRASDVPLGPNMRLAASEITLNDGFEVLHDHGPDLTLNRQIGSSVFRVETQBDGKH 118
Db 61 VDSR--DVPGLPEHPLRFATSETCLHGGFEVLHDHGPDLTLNRQIGSSVFRVEQPDGTH 118
QY 119 IAVGORNQGVTSVVLSDQEVARLQSIDPEGKDKFVFTGGRGGAGHAMVTVASDITEAR 178
Db 119 AAGVKDGVSVTVLNSSELQSLQSLDTEGNGRFVFTGGRGGGAGHAMVTVASDISAREK 178
QY 179 ILELLEPKGTG-----ESKGAGESKGVGELRESNGAENTTETOTSTSTSSLRSDP 229
Db 179 IIAKLDPDNHGGRQPKDITRSVGVGSASGMGD-----GV--VSETHTSTTSSVRS 230
QY 230 KWLALGTATVATGLIGLAATGIVQALALTPEDSPSTTTDPDAASATATATRDQLTKEA 289
Db 231 KFWVSGAIAAGLAGLAATGIVQALALTPEDPTTTDPDAANAASATKDQLTKEA 290
QY 290 NPDNOKVNIDELGNAPSGVLKDDVVANIEEOAKAAGEAKQQAIAENNAOAKKYDEQQA 349
Db 291 NPDNOKVNIDELGNAPSGELKDDVVQIAQQAQKEAGEVARQQAQVESNAQAQORHDDQA 350
QY 350 KRQBELKVSSGAGYGLSGALILGGGIGVAVTAALHRKNQPVQETTTTTTTTTSARTVE 409
Db 351 KRQBELDLSSGIGYGLSSALIVGGGIGAGVTAALHRRNPPTTEQTATTHS-----VIQ 403
QY 410 NKPAANTPAQGNVDTPGSEDTMESRRSSMASTSTSTPTDTSIGTVQNPYADV----- 460
Db 404 QQTGGNTPAQGGTDAETRAEDASLNRDSQGSVASTQWSDTS--GDVNPYAEGWMSRNP 461
QY 461 ----VKTSLHDSQVPTSNSTSVQNMGTDSVVYSTIOHPRPDITDNGARLGNPSAGIQ 516
Db 462 SLLAQPEEPIYDEVAADPN-----YSVQHFSGNPNVPTG--RLVSGPQGGIQ 505
QY 517 STYARLALSGGLRHDHMGGLTGGSNSAVNTSNPPAPGSHRFV 558
Db 506 STYALLANSGGLRLGMLGGLTGGSNSAGSAANAATPGVRFV 547

RESULT 11
Q9WXK1
ID AC Q9WXK1 PRELIMINARY; PRT; 547 AA.
DT 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Translocated intimin receptor.
GN Name=tir;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MPEC;
RA Okutani A., Itoh K., Sasakawa C.;
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB026719; BAA77400.1; -.
DR HSSP; Q9KWH9; 1FO2.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003536; Tir_receptor.
DR Pfam; PF07489; Tir_receptor C; 1.
DR Pfam; PF03549; Tir_receptor M; 1.
DR Pfam; PF07490; Tir_receptor N; 1.
DR PRINTS; PR01370; TRNSINTMINR.
KW Receptor.
SQ SEQUENCE 547 AA; 56283 MW; 02CAC6D625FA6EE1 CRC64;

```

```
Query Match      55.3%; Score 1569.5; DB 2; Length 547;
Best Local Similarity 57.0%; Pred. No. 4.5e-75;
Matches 332; Conservative 69; Mismatches 122; Indels 59; Gaps 11;

QY 1 MPIGNLGNPNVNNIIPAPPPLPSQTDGA--GGRQLINSTGPGSRALFTPPVRNMSADS 58
DB 1 MPIGNLGNNNISNNLIIPAPPPLPSQTDGATRGNGSSLISSGSLGSRLLFSLRSSIVDT 60

QY 59 GDNRASDPGLPVNPMRLAASEITLNDGFVLDHGGPLDTLNRQIGSSVFRVETQEDGKH 118
DB 61 VDQR--DVPGLEPEHPLRFATSETCLHGGFEVLHDKGPLDTLNRQIGSSVFRVETQEDGKH 118

QY 119 IAVGORNQVETSVLSDOEYARLQSIDPEGKDFVFTGGRGAGHAMVTVASDITEARQR 178
DB 119 AAVGKDGVEVSVTLNSELQSLDTEGNGRFTVFTGGRGSGHAMVTVASDISQAREK 178

QY 179 ILELLEPKGTG-----ESKGAGESKGVGELRESNGAENTTTTQTSTSTSLRSDP 229
DB 179 IIAKLDPNHHGGRQPKDIDTRSVGVSASGMD-----GV--VSETHSTTTTSSVRSDP 230

QY 230 KWLALGTVATGLIGLAATGIQVQALALTPEDSPPTTTPDAAASATETATRDQLTKEAFQ 289
DB 231 KFWVSVGAIAAGLAGLAATGIQVQALTPAPDPPTTTPDDEANAANAATAATKDQLTKEAFQ 290

QY 290 NPDNQKVNIDELGNAIPSGVLKDDVVAIEQAAGBEAKQQAENNAQAQKQYDEQQA 349
DB 291 NPDNQKVNIDELGNAIPSGELKDDVVAIQADQAKVAGEQQAQVESAQAQRHDDQQA 350

QY 350 KRQELKVSSGAGYGLSGALILGGIGVAVTAALHRKNQPVQETTTTTTTTTTSARTVE 409
DB 351 KRQELDLSSGIGYGLSSALIVGGIGAGVTAMLHRRNPTEQTIATTHS-----VIQ 403

QY 410 NKPNANTPAQGNVTPGSEDTHMESRRSSMASTST--PFDTSIGTVQNPYAD----- 460
DB 404 QQTGNTRAQGGADTTGVENASLRRDSQASVASTQMSDTS--GDVVPYAEQWMSRNNP 461

QY 461 ----VKTSLHDSQVPTSNSTSVQNMGNNTDSVVYSTIOHPPRDITDNGARLLGNPSAGIQ 516
DB 462 SLLAPEPIYDEVPDPN-----YSVIOHFGSNPNPTG-RLVGSFGQGIQ 505

QY 517 STYARLALSGGLRHDHMGGLTGGNSAVNTSNPPAPGSHRFV 558
DB 506 STYALLASSGGLRLGMGLTGGESAGSAANAATTPGVERFV 547

RESULT 12
Q9ETI1 PRELIMINARY; PRT; 547 AA.
AC Q9ETI1;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Translocated intimin receptor Tir.
OS Citrobacter rodentium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Citrobacter.
OX NCBI_TaxID=67825;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=CDC 1843-73T, and DBS100;
RX MEDLINE=20553330; PubMed=1101562;
RA Luperchio S.A., Newman J.V., Dangler C.A., Schrenzel M.D.,
RA Brenner D.J., Steigerwalt A.G., Schauer D.B.;
RT "Citrobacter rodentium, the causative agent of transmissible murine
RT colonic hyperplasia, exhibits clonality: synonymy of C. rodentium and
RT mouse-pathogenic Escherichia coli.";
RL J. Clin. Microbiol. 38:4343-4350 (2000).
DR EMBL; AF301618; AAG40758.1; -
DR EMBL; AF301617; AAG25642.1; -
DR HSSP; Q9KH9; 1F02.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
```

```
DR InterPro; IPR003536; Tir_receptor.
DR Pfam; PF07489; Tir_receptor C; 1.
DR Pfam; PF03549; Tir_receptor M; 1.
DR Pfam; PF07490; Tir_receptor N; 1.
DR PRINTS; PR01370; TRNSINTIMNR.
KW Receptor.
SQ SEQUENCE 547 AA; 56270 MW; CB8318B301049C37 CRC64;

Query Match      55.2%; Score 1568.5; DB 2; Length 547;
Best Local Similarity 57.0%; Pred. No. 5.1e-75;
Matches 332; Conservative 69; Mismatches 122; Indels 59; Gaps 11;

QY 1 MPIGNLGNPNVNNIIPAPPPLPSQTDGA--GGRQLINSTGPGSRALFTPPVRNMSADS 58
DB 1 MPIGNLGNNNISNNLIIPAPPPLPSQTDGATRGNGSSLISSGSLGSRLLFSLRSSIVDT 60

QY 59 GDNRASDPGLPVNPMRLAASEITLNDGFVLDHGGPLDTLNRQIGSSVFRVETQEDGKH 118
DB 61 VDQR--DVPGLEPEHPLRFATSETCLHGGFEVLHDKGPLDTLNRQIGSSVFRVETQEDGKH 118

QY 119 IAVGORNQVETSVLSDOEYARLQSIDPEGKDFVFTGGRGAGHAMVTVASDITEARQR 178
DB 119 AAVGKDGVEVSVTLNSELQSLDTEGNGRFTVFTGGRGSGHAMVTVASDISQAREK 178

QY 179 ILELLEPKGTG-----ESKGAGESKGVGELRESNGAENTTTTQTSTSTSLRSDP 229
DB 179 IIAKLDPNHHGGRQPKDIDTRSVGVSASGMD-----GV--VSETHSTTTTSSVRSDP 230

QY 230 KWLALGTVATGLIGLAATGIQVQALALTPEDSPPTTTPDAAASATETATRDQLTKEAFQ 289
DB 231 KFWVSVGAIAAGLAGLAATGIQVQALTPAPDPPTTTPDDEANAANAATAATKDQLTKEAFQ 290

QY 290 NPDNQKVNIDELGNAIPSGVLKDDVVAIEQAAGBEAKQQAENNAQAQKQYDEQQA 349
DB 291 NPDNQKVNIDELGNAIPSGELKDDVVAIQADQAKVAGEQQAQVESAQAQRHDDQQA 350

QY 350 KRQELKVSSGAGYGLSGALILGGIGVAVTAALHRKNQPVQETTTTTTTTTTSARTVE 409
DB 351 KRQELDLSSGIGYGLSSALIVGGIGAGVTAMLHRRNPTEQTIATTHS-----VIQ 403

QY 410 NKPNANTPAQGNVTPGSEDTHMESRRSSMASTST--PFDTSIGTVQNPYAD----- 460
DB 404 QQTGNTRAQGGADTTGVENASLRRDSQASVASTQMSDTS--GDVVPYAEQWMSRNNP 461

QY 461 ----VKTSLHDSQVPTSNSTSVQNMGNNTDSVVYSTIOHPPRDITDNGARLLGNPSAGIQ 516
DB 462 SLLAPEPIYDEVPDPN-----YSVIOHFGSNPNPTG-RLVGSFGQGIQ 505

QY 517 STYARLALSGGLRHDHMGGLTGGNSAVNTSNPPAPGSHRFV 558
DB 506 STYALLASSGGLRLGMGLTGGESAGSAANAATTPGVERFV 547

RESULT 13
O50190 PRELIMINARY; PRT; 549 AA.
AC O50190;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Translocated intimin receptor.
GN Name=tir;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=E2348/69;
RX MEDLINE=98050926; PubMed=9390560; DOI=10.1016/S0092-8674(00)80437-7;
RA Kenny B., Devinney R., Stein M., Reinscheid D.J., Frey E.A.,
RA Finlay B.B.;
RT "Enteropathogenic E. coli (EPEC) transfers its receptor for intimate
```

```

RT adherence into mammalian cells.;;
RN Cell 91:511-520(1997).
RA [2]
RP SEQUENCE FROM N.A.
RC STRAIN=E2348/69;
RA Stein M.S., Kenny B., Finlay B.B.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL, AF013122; AAB88410.1; -.
DR HSP; Q9KWH9; IF02.
DR GO: 0004872; F:receptor activity; IEA.
DR GO: 0007155; P:cell adhesion; IEA.
DR InterPro: IPR003536; TIR:receptor.
DR Pfam: PF07489; TIR:receptor_C; 1.
DR Pfam: PF03549; TIR:receptor_M; 1.
DR Pfam: PF07490; TIR:receptor_N; 1.
DR PRINTS; PR01370; TRNSINTMINR.
KW Receptor.
SQ SEQUENCE 549 AA; 56843 MW; 40C8B8B234409A08 CRC64;

Query Match 53.9%; Score 1531.5; DB 2; Length 549;
Best Local Similarity 56.2%; Pred. No. 4.6e-73;
Matches 327; Conservative 61; Mismatches 137; Indels 57; Gaps 11;

QY 1 MPIGNLGNPNVNSIPPAPLPSTQDGA--GGRGOLINSTGPGSLRALTTPVRNMSADS 58
DB 1 MPIGNLGNVNGNHLIPPAPLPSTQDGAARGTGHLISSTGALGSRSLFSPERNMSADS 60
QY 59 GDNRASDVPLGVNPMRLAA--SEITLNDGFVLDHDPGLDILNRQIGSSVFRVETQEDG 116
DB 61 VDSR--DIPGLPTNPSRLAATSETCLLGGFEVLHDKGPLDILNTQIGSPAFRVEQADG 118
QY 117 KHIAVGQNGVETSVLSQOEVARLOSIDPEKDKFVFTGGRGGAGHAWTVASDITEAR 176
DB 119 THAATGEKNGLEVSITLSPQEWSSLSQIDTEGKNRFVFTGGRGGHPMTVVASDIAEAR 178
QY 177 QRILELLEPKGTG-----ESKAGESKGVGELRESNGAENTTETOTSTSSLSRS 227
DB 179 TKILAKLDPDNGGQPKDVTFRSVGVGSAGI-----DDGV--VSEHTSTTSSVRS 230
QY 228 DPKLMLALGTVATGLGLAATGIVQALALTPEPDSPTTTDPDAASATETATRDQLTKEA 287
DB 231 DPKFVSVGAIAAGLAGLAATGIAQALALTPEPDDPTTTDPDAANAESATKDQLTQEA 290
QY 288 FQNPQKNVNDLGNALPSGLVDVVANIEEQAAGEAKQAQAIENNAQAQKYDQ 347
DB 291 FKNPENQKNVIDANGNAIPSGELKDDIVEQIAQAQEAQEAQAVESNAQAQORYEDQ 350
QY 348 QAKROELKVSAGYGLSGALILGGGIGVAVTAALHRKNQPVQETTTTTTTTTTSART 407
DB 351 HARRQELQLSGIGYGLSALIVAGGIGAGVTALHRKNQPAEQTTTITTTT-----HT 403
QY 408 VENKANNTPAQGNVDTPGSEDTMESRRSSMASTSTFTDTSIGTVQNPYADV-----K 462
DB 404 VVQQTGGIPQHKVALMPQERRRFRDRDSQGSVASTHWSDS--SEVNPYAEVGGARNS 462
QY 463 TSLHD-----SQVTSNNTSVQNMGTDSVVYSTIQHPDRTDNGARLGNPSAGIQ 516
DB 463 LSAHQPEEHIYDEVAADPGYVIONFSGSPVT-----GRLIGTPQGGIQ 507
QY 517 STYARLALSGRLHDMGGLTGGSNAVNTSNPPAPGSHRFV 558
DB 508 STYALLANSGLRLGMGLTSGGETAVSSVNAAPTGPVRFV 549

RESULT 14
Q79C12 PRELIMINARY; PRT; 367 AA.
AC Q79C12;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Similar to the hypothetical 38.4kDa protein of REPEC 84/110/1; unknown function (Fragment).

OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=REPEC RDEC-1;
RA Krejany E.O.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U59503; AAB02941.1; -.
DR GO: 0004872; F:receptor activity; IEA.
DR GO: 0007155; P:cell adhesion; IEA.
DR InterPro: IPR003536; TIR:receptor.
DR Pfam: PF07489; TIR:receptor_C; 1.
DR Pfam: PF03549; TIR:receptor_M; 1.
DR Pfam: PF07490; TIR:receptor_N; 1.
DR PRINTS; PR01370; TRNSINTMINR.
KW Hypothetical protein.
FT NON TER
SQ SEQUENCE 367 AA; 37797 MW; 00F84FCAC8FB5718 CRC64;

Query Match 40.4%; Score 1146.5; DB 2; Length 367;
Best Local Similarity 61.7%; Pred. No. 6.3e-53;
Matches 242; Conservative 38; Mismatches 83; Indels 29; Gaps 5;

QY 171 DITEARQRIILELLEPKGTGSGKAGESKGVGELRESNGAENTTETOTSTSSLSRSDPK 230
DB 1 DIAEARQRIIDKLEPKDKTKETKEPG-----DPNKGEGKIIHTSTSTSSLRADPK 51
QY 231 LWLALGTVATGLGLAATGIVQALALTPEPDSPTTTDPDAASATETATRDQLTKEAFQN 290
DB 52 LMLS LGTTAAGLIGNAATGIAQVALTPEPDDPTTTDPDAANTAEAAAKDQLTKEAFQN 111
QY 291 PDNQKNVIDELGNALPSGLVDVVANIEEQAAGEAKQAQAIENNAQAQKYDQ 350
DB 112 PDNQKNVIDENGNALPSGELKDDVVQAIAEQAKAAGEAQARQAIENSQAQKYDEQHAH 171
QY 351 ROELKVSAGYGLSGALILGGGIGVAVTAALHRKNQPVQETTTTTTTTTTSARTVEN 410
DB 172 REQEMLSGGYGLSGALILGGGIGAGVTALHRKNQPAEQTTTITTTT-----VVDN 223
QY 411 KPANTPAQGNVDTPGSEDTMESRRSSMASTSTFTDTSIGTVQNPYADVKTSLHDSQV 470
DB 224 QTNNAAGQNTDTSGPRESPASRRNSNASLASGDSSTSTGTVENPYADV-----CM 276
QY 471 PTSNNTSVQNMGTDSVV-----YSTIQHPDRTDNGARLGNPSAGIQSTYARLALSG 526
DB 277 PRNDSLARISEPIYDEVAADPNYSVIOHFGSNSPVTG-RLVGTGPGQIGSTVALLASG 335
QY 527 GLRHDGGLTGGSNAVNTSNPPAPGSHRFV 558
DB 336 GLRLGMGLTGGSNAVSTANAAPTGPVRFV 367

RESULT 15
Q7YU77 PRELIMINARY; PRT; 1374 AA.
AC Q7YU77;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-WAR-2004 (TrEMBLrel. 26, Last annotation update)
DE SD04152p.
GN Name=Alhambra;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champs M., Chavez C., Dorsett V., Dreenek D., Farfan D., Friese E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,

```

RA	3	IGNLGNPNVANSIPPA	PSQTDGAGRGQLINSTGPLGSRALFTPV	VRNSMADSGDNR	62	
RA	375	IGNISN--SLNNLP	CG--SSSTSSAG-----NVP	GSGGIIISAGSGATOSTSQ	421	
RA	63	ASDVPGLVPNMR	LAASEITLNDGFEVLHDHGPLD	TLNRIGQSSVFRVETQBDG	KHIAVG 122	
RA	422	SSTAPGTTKSS---	ASSSSSSNSYKEKHS---KSL	SKSTSSK-----DKDGD	SSNT 468	
RA	123	QRNGVET----	SVLSLDOEYARLQSID	PEGDK-----VFV	154	
RA	469	SANNFTNSSAS	TSNSSSTREKSSSKLKNKDSNQ	VPSATSSLTSSINTQ	SSSTA 528	
RA	155	TGGRGGAG-HAMVT	VASDITEARQRIE-----LLE	PKGTGESKAGESKGVGL	RESN 207	
RA	529	TAGSGGTGTHV	SSSSAAGLSNAPSTTNEH	NHNLSTNGTGAGSAGK	LOSVNLSSS 588	
RA	208	SG-----	AEHTTETOTSTSSLR	DPKLWALGTVATGLIGLAA--	TGIVQAL 254	
RA	589	SGFGSLRSV	TSSSSTVNDSTGGFGNS	SNERENLSGAGSSASN	MPGTIAPGTGCVSSS 648	
RA	255	ALTEPDSPTT	DDPAAASATETATRDQ	LTKFAFQNPDKNVIDEL	GNAIPSGVLKDV 314	
RA	649	AATNLSTNKG	SSSSTANSLTSTSTP-----	SGSSSSSSSKRRKAD-----	689	
RA	315	VANIEEQA	KAGE-EAKQQAENNAQA	KKYD-----EQQAKRQE-E	354	
RA	690	-----	SAKSTSSITSGSALED	NNLSIRYDIKDQVQAL	PLTDFEKEIKSSKRQTE 743	
RA	355	LKV---	SSGAGYGLSGALILGG	GIGVAVTAALHRKNQ	PVEQTTTTTTTTTTSARTVENK 411	
RA	744	LSPPTHOTS	ATAEVNAPLASSTASIA	VTASATAASAPPATCT	TTTILGSSISISGNAGSTSG 803	
RA	412	PANNTPAQ	GNVDT-----PGSEDT	MESRRSSMASTSTSTP	PDSS-----IGTVQ 455	
RA	804	NSSGV	SATGGAQSAVSGSGY	PKTESKSGTASAGSG	SSSNTSTSKHGSNIKDISSS 863	
RA	456	NPYADVKT	-----SLHDSQ	VPSTNSNTSVQNN	MGNTDSVVYSTIOHP	PRDPTDONGAR---- 506
RA	864	NOQASTASS	SSSAPSLYVS-VPL	STANVPGLNPLTST	SSSTTSSSHSASRRSG	GAQSQHQ 922
RA	507	-----	LLGNPSAGIQSTY	ARLALSGGLRHD	MGGLTGGSNSAV-----	NTSNNPPAGS 554
RA	923	QQLSNALV	GPMPMG-----	SAGAFH--GGTTS	AGSSSVIQQSGKSP	PALGT 967

Search completed: May 13, 2005, 11:47:35
Job time : 182 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 13, 2005, 11:57:23 ; Search time 42 Seconds
(without alignments)
1278.308 Million cell updates/sec

Title: US-09-189-415b-11
Perfect score: 558
Sequence: 1 MPIGNLGNPNVNSIPPAP.....SNSAVNTSNPPAGSHRFV 558

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 8

Total number of hits satisfying chosen parameters: 210

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : PIR 79:*

1: Piri:*

2: Piri:*

3: Piri:*

4: Piri:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	558	100.0	558	2 A98199	translocated intim
2	558	100.0	558	2 E86045	probable transloca
3	13	2.3	166	2 C90029	hypothetical prote
4	12	2.2	139	2 D86417	probable auxin-ind
5	12	2.2	458	2 T31631	hypothetical prote
6	12	2.2	524	2 S33640	homeotic protein s
7	12	2.2	560	2 T32661	hypothetical prote
8	12	2.2	569	2 S47277	gp88 protein - mur
9	12	2.2	802	2 A36910	xylanase, beta(1,3
10	12	2.2	1023	2 S12519	glutactin - fruit
11	12	2.2	3712	2 S18253	laminin alpha-1 ch
12	12	2.2	4377	2 A55575	ankyrin 3, long sp
13	11	2.0	67	2 B56888	alkaline phosphata
14	11	2.0	108	2 T26880	hypothetical prote
15	11	2.0	164	2 T26561	hypothetical prote
16	11	2.0	183	2 S03358	hypothetical prote
17	11	2.0	208	2 T46896	merozoite surface
18	11	2.0	217	2 S01358	salivary glue prot
19	11	2.0	234	2 T25660	hypothetical prote
20	11	2.0	245	2 T26868	hypothetical prote
21	11	2.0	274	2 A45632	merozoite surface
22	11	2.0	278	2 S39310	merozoite surface
23	11	2.0	284	2 T22023	hypothetical prote
24	11	2.0	304	2 T15922	hypothetical prote
25	11	2.0	327	2 S20074	promastigote surfa
26	11	2.0	341	2 T32949	hypothetical prote
27	11	2.0	342	2 T29557	hypothetical prote
28	11	2.0	373	2 T29596	hypothetical prote
29	11	2.0	385	2 JC7783	RAD 23B protein -

30	11	2.0	415	2 T32467	hypothetical prote
31	11	2.0	468	2 A5476	protein kinase (EC
32	11	2.0	477	2 A54843	nemo, form 1 - fru
33	11	2.0	484	2 S58868	G protein-coupled
34	11	2.0	512	2 T02498	probable WRKY-type
35	11	2.0	516	2 S19252	1-aminocyclopropan
36	11	2.0	517	2 T30658	probable zinc meta
37	11	2.0	518	2 S31442	1-aminocyclopropan
38	11	2.0	519	2 T37339	hypothetical prote
39	11	2.0	525	2 A35596	nuclear pore glyco
40	11	2.0	526	2 A56573	nuclear pore compl
41	11	2.0	530	2 T32812	hypothetical prote
42	11	2.0	551	2 S18408	alkaline phosphata
43	11	2.0	559	2 B36307	alkaline phosphata
44	11	2.0	651	2 T21175	hypothetical prote
45	11	2.0	680	2 T19939	hypothetical prote
46	11	2.0	681	2 T23454	hypothetical prote
47	11	2.0	698	2 A54796	regulatory protein
48	11	2.0	732	2 T25937	hypothetical prote
49	11	2.0	770	2 T22808	hypothetical prote
50	11	2.0	781	2 S51592	xyNB precursor - R
51	11	2.0	825	2 T29634	hypothetical prote
52	11	2.0	831	2 T08611	hypothetical prote
53	11	2.0	889	2 A35679	rep protein - slim
54	11	2.0	975	2 T08606	protein phosphatas
55	11	2.0	1002	2 T30546	major surface glyc
56	11	2.0	1076	2 JC2217	major surface glyc
57	11	2.0	1083	2 JC2300	cell surface glyco
58	11	2.0	1089	2 T14576	nosa protein - eli
59	11	2.0	1093	2 T18275	1-phosphatidylinos
60	11	2.0	1099	2 T18257	phospholipase C -
61	11	2.0	1271	2 D84237	hypothetical prote
62	11	2.0	1282	2 JE0120	glycoprotein A - m
63	11	2.0	1402	2 T17456	cell surface prote
64	11	2.0	1635	2 T14075	chitinase (EC 3.2.
65	11	2.0	1671	2 S71628	sensory transducti
66	11	2.0	1737	2 A59235	unconventional myo
67	11	2.0	1832	2 T31113	mucin-like glycopo
68	11	2.0	1858	2 T18273	1-phosphatidylinos
69	10	1.8	127	2 T51538	nitrilase associat
70	10	1.8	187	2 T49491	hypothetical prote
71	10	1.8	195	2 T19617	hypothetical prote
72	10	1.8	263	2 S01360	salivary glue prot
73	10	1.8	307	1 GSFF3	salivary glue prot
74	10	1.8	388	2 T16861	hypothetical prote
75	10	1.8	390	2 T49619	hypothetical prote
76	10	1.8	393	2 B86189	protein T25N20.9 l
77	10	1.8	395	2 T45599	hypothetical prote
78	10	1.8	435	2 T25350	hypothetical prote
79	10	1.8	572	2 T16865	hypothetical prote
80	10	1.8	577	2 G89430	protein K02E2.3 (l
81	10	1.8	648	1 JQ1150	protein kinase (EC
82	10	1.8	649	2 T24505	hypothetical prote
83	10	1.8	662	2 A45155	mucin FIM-C.1 - Af
84	10	1.8	788	2 S05661	muscarinic acetyl
85	10	1.8	816	2 C69493	hypothetical prote
86	10	1.8	947	2 T08605	hypothetical prote
87	10	1.8	1008	2 T30544	major surface glyco
88	10	1.8	1014	2 T18759	hypothetical prote
89	10	1.8	1017	2 T30542	major surface glyco
90	10	1.8	1022	2 T30543	major surface glyco
91	10	1.8	1030	2 T18374	B-cell receptor pr
92	10	1.8	1390	2 T14004	trfA protein - eli
93	10	1.8	1513	2 T23681	hypothetical prote
94	10	1.8	1570	2 T18272	1-phosphatidylinos
95	9	1.6	124	2 T48833	hypothetical prote
96	9	1.6	202	2 F8755	prophage p12 prote
97	9	1.6	213	2 T23865	hypothetical prote
98	9	1.6	224	2 G86148	TiN6.14 protein -
99	9	1.6	232	2 A60095	larval glue protei
100	9	1.6	492	2 A41907	methyl-CpG-binding
101	9	1.6	500	1 EPFF	zip protein precur
102	9	1.6	622	2 T51223	hypothetical prote

103	9	1.6	645	2	T29818	176	8	1.4	730	2	A53064	folded gastrulation
104	9	1.6	708	2	T29669	177	8	1.4	771	2	T21633	hypothetical prote
105	9	1.6	906	2	S03313	178	8	1.4	786	1	A47547	serine proteinase
106	9	1.6	943	2	S68824	179	8	1.4	790	2	T34293	hypothetical prote
107	9	1.6	1408	2	S16148	180	8	1.4	802	2	T21315	hypothetical prote
108	9	1.6	1510	2	T33100	181	8	1.4	816	2	S84439	hypothetical prote
109	9	1.6	3672	2	T23433	182	8	1.4	854	2	T03107	glycoprotein B - a
110	9	1.6	3704	2	T37316	183	8	1.4	892	2	T06818	DNA topoisomerase
111	8	1.4	76	2	B96809	184	8	1.4	907	2	A45560	sporozoite surface
112	8	1.4	78	2	A75442	185	8	1.4	918	2	D71407	hypothetical prote
113	8	1.4	80	2	A91255	186	8	1.4	936	2	T06190	lipoxigenase (EC 1
114	8	1.4	80	2	E86095	187	8	1.4	977	2	T16232	hypothetical prote
115	8	1.4	80	2	A65210	188	8	1.4	1013	2	T33470	hypothetical prote
116	8	1.4	81	2	AE1013	189	8	1.4	1047	2	A55617	masquerade precurs
117	8	1.4	83	2	T20306	190	8	1.4	1057	2	T04874	hypothetical prote
118	8	1.4	105	2	T06085	191	8	1.4	1096	2	T08619	probable metallopr
119	8	1.4	147	2	T01039	192	8	1.4	1129	2	T25635	hypothetical prote
120	8	1.4	167	2	T33602	193	8	1.4	1172	2	T00065	hypothetical prote
121	8	1.4	169	2	T32698	194	8	1.4	1175	2	T25634	hypothetical prote
122	8	1.4	173	2	S16162	195	8	1.4	1229	2	T25697	hypothetical prote
123	8	1.4	176	2	T26212	196	8	1.4	1269	2	T14476	p51A protein - eli
124	8	1.4	180	2	B45613	197	8	1.4	1272	2	T30248	fragile X mental r
125	8	1.4	193	2	T24370	198	8	1.4	1279	2	T13613	hypothetical prote
126	8	1.4	210	2	T49785	199	8	1.4	1314	1	TNBYR6	transcription regu
127	8	1.4	225	2	S59852	200	8	1.4	1335	2	T18289	racGAP protein - s
128	8	1.4	240	2	AD3584	201	8	1.4	1409	2	T29594	hypothetical prote
129	8	1.4	270	2	T41759	202	8	1.4	1603	2	S17983	gene posterior sex
130	8	1.4	272	2	G71618	203	8	1.4	1813	2	T30564	resistance protein
131	8	1.4	275	2	S09774	204	8	1.4	1969	2	T08875	histidine kinase h
132	8	1.4	286	2	B45632	205	8	1.4	3020	2	A43932	mucin 2 precursor,
133	8	1.4	287	2	B39615	206	8	1.4	3133	2	S52093	hemocytin - silkwo
134	8	1.4	300	2	A39112	207	8	1.4	3216	2	C90538	hypothetical prote
135	8	1.4	302	2	A39615	208	8	1.4	3828	2	T13857	trithorax protein
136	8	1.4	302	2	C86480	209	8	1.4	3848	2	T17414	TipC protein - eli
137	8	1.4	303	2	T03571	210	8	1.4	5170	2	T15348	hypothetical prote
138	8	1.4	314	2	D96703							
139	8	1.4	327	2	T49514							
140	8	1.4	333	2	T45720							
141	8	1.4	333	2	F84775							
142	8	1.4	347	2	B39112							
143	8	1.4	381	2	T09640							
144	8	1.4	384	2	A44146							
145	8	1.4	389	2	T33340							
146	8	1.4	390	2	T02845							
147	8	1.4	394	2	T20633							
148	8	1.4	406	2	T16520							
149	8	1.4	416	2	T14021							
150	8	1.4	422	2	T49513							
151	8	1.4	434	2	T47772							
152	8	1.4	444	2	C87297							
153	8	1.4	467	2	A60667							
154	8	1.4	467	2	A45629							
155	8	1.4	475	2	T27811							
156	8	1.4	490	2	T02843							
157	8	1.4	496	2	E90181							
158	8	1.4	513	2	F83162							
159	8	1.4	531	2	S57752							
160	8	1.4	540	2	AD0522							
161	8	1.4	557	2	H69678							
162	8	1.4	558	2	JC2004							
163	8	1.4	587	2	D84426							
164	8	1.4	591	2	T25636							
165	8	1.4	592	2	A32107							
166	8	1.4	592	2	T34446							
167	8	1.4	621	2	T20307							
168	8	1.4	634	2	A64521							
169	8	1.4	639	2	G02919							
170	8	1.4	668	2	A71986							
171	8	1.4	672	2	T20310							
172	8	1.4	672	2	T32557							
173	8	1.4	710	2	A99486							
174	8	1.4	712	1	I46031							
175	8	1.4	721	2	B90487							


```
Db 121 VQQRNGVTSVLSQEVYARLQSIIDPEGKDKVFVFGGRGAGHAMVTVASDITEARQIL 180
QY 181 ELLEPKGTGSKGAGESKGVGLRESNSGAENTTETQTSTSSLRSDPKLMLGTVAT 240
Db 181 ELLEPKGTGSKGAGESKGVGLRESNSGAENTTETQTSTSSLRSDPKLMLGTVAT 240
QY 241 GLIGLAATGIVQALALTTPEDSPPTTDPDAASATETATRDQLTKEAFQNPQKNWIDE 300
Db 241 GLIGLAATGIVQALALTTPEDSPPTTDPDAASATETATRDQLTKEAFQNPQKNWIDE 300
QY 301 LGNAIPSGVLKDDVVANIEEQAKAAGEAKQAIIENNAQAQKKYDEQQAQKEELKVSSG 360
Db 301 LGNAIPSGVLKDDVVANIEEQAKAAGEAKQAIIENNAQAQKKYDEQQAQKEELKVSSG 360
QY 361 AGYGLSGALILGGGIGVAVTAALHRKNQPVETTTTTTTTTTTTTTSARTVENKPNNTPAQG 420
Db 361 AGYGLSGALILGGGIGVAVTAALHRKNQPVETTTTTTTTTTTTTTSARTVENKPNNTPAQG 420
QY 421 NVDTPGSEDTMESRSSMASTSTSTFDTSIGTVQNPYADVKTSLHDSOVPTSNSNTSVQ 480
Db 421 NVDTPGSEDTMESRSSMASTSTSTFDTSIGTVQNPYADVKTSLHDSOVPTSNSNTSVQ 480
QY 481 NMGNNTDSVVYSTIQHPPRDTTNGARLLGNPSAGIQSTYARLALSGGLRHDGGLTGGSN 540
Db 481 NMGNNTDSVVYSTIQHPPRDTTNGARLLGNPSAGIQSTYARLALSGGLRHDGGLTGGSN 540
QY 541 SAVNTSNNPPAPGSHRFV 558
Db 541 SAVNTSNNPPAPGSHRFV 558

RESULT 2
E86045
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: E86045
R:Perina, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glaesner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grobbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: E86045
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-558 <STO>
A:Cross-references: UNIPROT:O9R396; GB:AE005174; NID:G12518449; PIDN:AA058825.1; GSPDB:G
A:Experimental source: strain O157:H7, substrain EDL933
A:Gene: tir

Query Match 100.0%; Score 558; DB 2; Length 558;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 558; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPIGNLGHNPVNNISIPAPPSPQTDGAGRGQLINSTGPGSRALFPPVNSMADSGD 60
Db 1 MPIGNLGHNPVNNISIPAPPSPQTDGAGRGQLINSTGPGSRALFPPVNSMADSGD 60

QY 61 NRASDVPGLPVNPMLAASEITLNDGFVLDHGPDLTLNRQIGSSVFVETQEDGKHIA 120
Db 61 NRASDVPGLPVNPMLAASEITLNDGFVLDHGPDLTLNRQIGSSVFVETQEDGKHIA 120

QY 121 VQQRNGVTSVLSQEVYARLQSIIDPEGKDKVFVFGGRGAGHAMVTVASDITEARQIL 180
Db 121 VQQRNGVTSVLSQEVYARLQSIIDPEGKDKVFVFGGRGAGHAMVTVASDITEARQIL 180
QY 181 ELLEPKGTGSKGAGESKGVGLRESNSGAENTTETQTSTSSLRSDPKLMLGTVAT 240
Db 181 ELLEPKGTGSKGAGESKGVGLRESNSGAENTTETQTSTSSLRSDPKLMLGTVAT 240
QY 241 GLIGLAATGIVQALALTTPEDSPPTTDPDAASATETATRDQLTKEAFQNPQKNWIDE 300
```

```
Db 241 GLIGLAATGIVQALALTTPEDSPPTTDPDAASATETATRDQLTKEAFQNPQKNWIDE 300
QY 301 LGNAIPSGVLKDDVVANIEEQAKAAGEAKQAIIENNAQAQKKYDEQQAQKEELKVSSG 360
Db 301 LGNAIPSGVLKDDVVANIEEQAKAAGEAKQAIIENNAQAQKKYDEQQAQKEELKVSSG 360
QY 361 AGYGLSGALILGGGIGVAVTAALHRKNQPVETTTTTTTTTTTTTTSARTVENKPNNTPAQG 420
Db 361 AGYGLSGALILGGGIGVAVTAALHRKNQPVETTTTTTTTTTTTTTSARTVENKPNNTPAQG 420
QY 421 NVDTPGSEDTMESRSSMASTSTSTFDTSIGTVQNPYADVKTSLHDSOVPTSNSNTSVQ 480
Db 421 NVDTPGSEDTMESRSSMASTSTSTFDTSIGTVQNPYADVKTSLHDSOVPTSNSNTSVQ 480
QY 481 NMGNNTDSVVYSTIQHPPRDTTNGARLLGNPSAGIQSTYARLALSGGLRHDGGLTGGSN 540
Db 481 NMGNNTDSVVYSTIQHPPRDTTNGARLLGNPSAGIQSTYARLALSGGLRHDGGLTGGSN 540
QY 541 SAVNTSNNPPAPGSHRFV 558
Db 541 SAVNTSNNPPAPGSHRFV 558

RESULT 3
C90029
hypothetical protein SA2097 [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: C90029
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kato, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A99758; MUID:213111952; PMID:11418146
A:Accession: C90029
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-166 <KUR>
A:Cross-references: UNIPROT:Q99RW9; GB:BA000018; PID:G13702104; PIDN:BA043396.1; GSPDB:G
A:Experimental source: strain N315
A:Genetics:
A:Gene: SA2097

Query Match 2.3%; Score 13; DB 2; Length 166;
Best Local Similarity 100.0%; Pred. No. 0.00016;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 392 QTTTITTTTTTTS 404
Db 43 QTTTITTTTTTTS 55

RESULT 4
D86417
probable auxin-induced protein, 50455-50036 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: D86417
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Hultzar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luo, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: D86417
A>Status: preliminary
```

```
A:Molecule type: DNA
A:Residues: 1-139 <STO>
A:Cross-references: UNIPROT:Q9C7Q5; GB:AE005172; NID:G10092232; PIDN:AGG12648.1; GSPDB:G
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T31631
R:Smyle, R.
submitted to the EMBL Data Library, September 1999
A:Reference number: Z21048
A:Accession: T31631
A:Status: preliminary; translated from GB/EMBL/DBBJ
A:Molecule type: DNA
A:Residues: 1-458 <WIL>
A:Cross-references: UNIPROT:Q9NA83; EMBL:AL117195; NID:el549729; PIDN:CAB55014.1; CESP:Y
C:Genetics:
A:Experimental source: clone Y57A10A
A:Gene: CESP:Y57A10A.1
A:Introns: 8/3; 54/3; 112/3; 151/1

Query Match      2.2%; Score 12; DB 2; Length 458;
Best Local Similarity 100.0%; Pred. No. 0.0037;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 393 TTTTITTTTTTTS 404
DB 135 TTTTITTTTTTTS 146

RESULT 6
S33640
homeotic protein smox-2, engrailed-like - fluke (Schistosoma mansoni)
C:Species: Schistosoma mansoni
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 16-Aug-2004
C:Accession: S33640; S27841
R:Webster, P.J.; Mansour, T.E.
Mech. Dev. 38, 25-32, 1992
A:Title: Conserved classes of homeodomains in Schistosoma mansoni, an early bilateral me
A:Reference number: S33640; MUID:92399260; PMID:1356008
A:Accession: S33640
A:Molecule type: mRNA
A:Residues: 1-524 <WEB>
A:Cross-references: UNIPROT:Q266601; EMBL:S441191; EMBL:M85305; NID:G161103; PIDN:AAA29929
C:Genetics:
A:Gene: smox-2
C:Superfamily: homeobox homology
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
F:424-480/Domain: homeobox homology <HOX>

Query Match      2.2%; Score 12; DB 2; Length 524;
Best Local Similarity 100.0%; Pred. No. 0.0042;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 393 TTTTITTTTTTTS 404
DB 105 TTTTITTTTTTTS 116

RESULT 7
T31631
hypothetical protein Y57A10A.i - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T31631
R:Smyle, R.
submitted to the EMBL Data Library, September 1999
A:Reference number: Z21048
A:Accession: T31631
A:Status: preliminary; translated from GB/EMBL/DBBJ
A:Molecule type: DNA
A:Residues: 1-458 <WIL>
A:Cross-references: UNIPROT:Q9NA83; EMBL:AL117195; NID:el549729; PIDN:CAB55014.1; CESP:Y
C:Genetics:
A:Experimental source: clone Y57A10A
A:Gene: CESP:Y57A10A.1
A:Introns: 8/3; 54/3; 112/3; 151/1

Query Match      2.2%; Score 12; DB 2; Length 139;
Best Local Similarity 100.0%; Pred. No. 0.0013;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 393 TTTTITTTTTTTS 404
DB 22 TTTTITTTTTTTS 33

RESULT 5
T31631
hypothetical protein Y57A10A.i - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T31631
R:Smyle, R.
submitted to the EMBL Data Library, September 1999
A:Reference number: Z21048
A:Accession: T31631
A:Status: preliminary; translated from GB/EMBL/DBBJ
A:Molecule type: DNA
A:Residues: 1-458 <WIL>
A:Cross-references: UNIPROT:Q9NA83; EMBL:AL117195; NID:el549729; PIDN:CAB55014.1; CESP:Y
C:Genetics:
A:Experimental source: clone Y57A10A
A:Gene: CESP:Y57A10A.1
A:Introns: 8/3; 54/3; 112/3; 151/1

Query Match      2.2%; Score 12; DB 2; Length 139;
Best Local Similarity 100.0%; Pred. No. 0.0013;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 393 TTTTITTTTTTTS 404
DB 22 TTTTITTTTTTTS 33

RESULT 5
T32661
hypothetical protein K11D12.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T32661
R:Henkhaus, J.; Wohldmann, P.; Gillam, B.
submitted to the EMBL Data Library, December 1997
A:Description: The sequence of C. elegans cosmid K11D12.
A:Reference number: Z21207
A:Accession: T32661
A:Status: preliminary; translated from GB/EMBL/DBBJ
A:Molecule type: DNA
A:Residues: 1-560 <HEN>
A:Cross-references: EMBL:AF039047; PIDN:AAB94223.1; GSPDB:GN00023; CESP:K11D12.1
A:Experimental source: strain Bristol N2; clone K11D12
C:Genetics:
A:Gene: CESP:K11D12.1
A:Map position: 5
A:Introns: 5/3; 48/3; 90/3; 127/3; 149/3; 190/1; 207/1; 233/3; 264/1; 480/1

Query Match      2.2%; Score 12; DB 2; Length 560;
Best Local Similarity 100.0%; Pred. No. 0.0044;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 393 TTTTITTTTTTTS 404
DB 348 TTTTITTTTTTTS 359

RESULT 8
S47277
gp88 protein - murine cytomegalovirus
C:Species: murine cytomegalovirus, murine herpesvirus 1
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C:Accession: S47277
R:Thaete, R.; Lucin, P.; Schneider, K.; Koszinowski, U.
submitted to the EMBL Data Library, February 1994
A:Reference number: S47277
A:Accession: S47277
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-569 <THA>
A:Cross-references: UNIPROT:Q83183; EMBL:X77798; NID:G535195; PIDN:CAAS4825.1; PID:G53519
C:Superfamily: murine cytomegalovirus gp88 protein

Query Match      2.2%; Score 12; DB 2; Length 569;
Best Local Similarity 100.0%; Pred. No. 0.0045;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 392 QTTTITTTTTTTT 403
DB 472 QTTTITTTTTTTT 483

RESULT 9
A36910
xylanase, beta(1,3-1,4)-glucanase - Ruminococcus flavefaciens
C:Species: Ruminococcus flavefaciens
C:Date: 07-Apr-1994 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: A36910
R:Flint, H.J.; Martin, J.; McPherson, C.A.; Daniel, A.S.; Zhang, J.X.
J. Bacteriol. 175, 2943-2951, 1993
A:Title: A bifunctional enzyme, with separate xylanase and beta(1,3-1,4)-glucanase domain
A:Reference number: A36910; MUID:93259938; PMID:8491715
A:Accession: A36910
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-802 <FLI>
A:Cross-references: UNIPROT:Q9S310; GB:S61204; NID:G385910; PIDN:AAB26620.1; PID:G385911
A:Note: sequence extracted from NCBI backbone (NCBIN:131871, NCBI:P:131872)
F:42-239/Domain: endo-1,4-beta-xylanase homology <XYL>
F:259-401/Domain: Thermotoga xylanase A amino-terminal repeat homology <TXA>
```

```
Query Match      2.2%; Score 12; DB 2; Length 802;
Best Local Similarity 100.0%; Pred. No. 0.0061;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 393 TTTTITTTTTTS 404
    |||||||
Db 533 TTTTITTTTTTS 544

RESULT 10
12519
glutactin - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C>Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: S12519
R:Olson, P.F.; Fessler, L.I.; Nelson, R.E.; Sterne, R.E.; Campbell, A.G.; Fessler, J.H.
EMBO J. 9, 1219-1227, 1990
A:Title: Glutactin, a novel Drosophila basement membrane-related glycoprotein with sequ
A:Reference number: S12519; MUID:90214632; PMID:2108864
A:Accession: S12519
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1023 <OLS>
A:Cross-references: UNIPROT:P33438; EMBL:X53286; NID:G297084; PIDN:CAA37380.1; PID:G2970
C:Genetics:
A:Introns: 390/3

Query Match      2.2%; Score 12; DB 2; Length 1023;
Best Local Similarity 100.0%; Pred. No. 0.0075;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 393 TTTTITTTTTTS 404
    |||||||
Db 605 TTTTITTTTTTS 616

RESULT 11
S18253
laminin alpha-1 chain precursor - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C>Date: 16-Sep-1992 #sequence_revision 24-Jul-1997 #text_change 09-Jul-2004
R:Garrison, K.; MacKrell, A.J.; Fessler, J.H.
C:Accession: S28399; S18253
R:Kusche-Gullberg, M.; Garrison, K.; MacKrell, A.J.; Fessler, L.I.; Fessler, J.H.
EMBO J. 11, 4519-4527, 1992
A:Title: Laminin A chain: expression during Drosophila development and genomic sequence.
A:Reference number: S28399; MUID:93049203; PMID:1425586
A:Accession: S28399
A>Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-3712 <KUS>
A:Cross-references: UNIPROT:Q00174; GB:M96388; NID:g157799; PIDN:AAA28662.1; PID:g157800
R:Garrison, K.; MacKrell, A.J.; Fessler, J.H.
J. Biol. Chem. 266, 22899-22904, 1991
A:Title: Drosophila laminin A chain sequence, interspecies comparison, and domain structu
A:Reference number: S18253; MUID:92078147; PMID:1744083
A:Accession: S18253
A:Molecule type: mRNA
A:Residues: 1762-3712 <GAR>
A:Cross-references: EMBL:M75882; NID:g157797; PIDN:AAA28661.1; PID:g157798
C:Genetics:
A:Gene: FlyBase:LaNa
A:Cross-references: FlyBase:FBgn0002526
C:Superfamily: laminin alpha-1 chain; laminin G repeat homology; laminin-type EGF-like H
C:Keywords: basement membrane; cell binding; coiled coil; disulfide bond; extracellular
F:273-330/Domain: laminin-type EGF-like homology <LEG>
F:333-400/Domain: laminin-type EGF-like homology <LE02>
F:541-584/Domain: laminin-type EGF-like homology <LEG1>
F:1776-2115/Domain: III <DOM3>
F:1776-1806/Domain: laminin-type EGF-like homology #status atypical <LE1>
F:1809-1856/Domain: laminin-type EGF-like homology <LE2>
F:1859-1914/Domain: laminin-type EGF-like homology <LE3>
F:1917-1967/Domain: laminin-type EGF-like homology <LE4>
F:1970-2014/Domain: laminin-type EGF-like homology <LE5>
```

```
F:2017-2061/Domain: laminin-type EGF-like homology <LE6>
F:2064-2109/Domain: laminin-type EGF-like homology <LE7>
F:2116-2697/Domain: I/II, heptad repeats <DOM2>
F:2698-3712/Domain: G <DOMG>
F:2698-2863/Domain: repeat G1 <RG1>
F:2864-3048/Domain: repeat G2 <RG2>
F:3049-3223/Domain: repeat G3 <RG3>
F:3079-3200/Domain: laminin G repeat homology <LG3>
F:3334-3528/Domain: repeat G4 <RG4>
F:3529-3712/Domain: repeat G5 <RG5>
F:1847,1850,1943,2024,2196,2215,2267,2301,2323,2482,2524,2538,2569,2699,2720,2890,2938,30
Query Match      2.2%; Score 12; DB 2; Length 3712;
Best Local Similarity 100.0%; Pred. No. 0.023;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 393 TTTTITTTTTTS 404
    |||||||
Db 3273 TTTTITTTTTTS 3284

RESULT 12
A55575
ankyrin 3, long splice form - human
N:Alternate names: ankyrin G
C:Species: Homo sapiens (man)
C>Date: 03-Mar-1995 #sequence_revision 03-Mar-1995 #text_change 09-Jul-2004
C:Accession: A55575
R:Kordeli, E.; Lambert, S.; Bennett, V.
J. Biol. Chem. 270, 2352-2359, 1995
A:Title: Ankyrin-G, A new ankyrin gene with neural-specific isoforms localized at the ax
A:Reference number: A55575; MUID:95138209; PMID:7836469
A:Accession: A55575
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-4377 <KOR>
A:Cross-references: UNIPROT:Q12955; GB:U13616; NID:G608024; PIDN:AAA64834.1; PID:G608025
C:Genetics:
A:Gene: GDB:ANK3
A:Cross-references: GDB:424503; OMIM:600465
A:Map position: 10q21-10q21
C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology
C:Keywords: alternative splicing; peripheral membrane protein
F:73-105/Domain: ankyrin repeat homology <AN01>
F:106-138/Domain: ankyrin repeat homology <AN02>
F:139-171/Domain: ankyrin repeat homology <AN03>
F:172-200/Domain: ankyrin repeat homology <AN04>
F:201-233/Domain: ankyrin repeat homology <AN05>
F:234-266/Domain: ankyrin repeat homology <AN06>
F:267-299/Domain: ankyrin repeat homology <AN07>
F:300-332/Domain: ankyrin repeat homology <AN08>
F:333-365/Domain: ankyrin repeat homology <AN09>
F:366-398/Domain: ankyrin repeat homology <AN10>
F:399-431/Domain: ankyrin repeat homology <AN11>
F:432-464/Domain: ankyrin repeat homology <AN12>
F:465-497/Domain: ankyrin repeat homology <AN13>
F:498-530/Domain: ankyrin repeat homology <AN14>
F:531-563/Domain: ankyrin repeat homology <AN15>
F:564-596/Domain: ankyrin repeat homology <AN16>
F:597-629/Domain: ankyrin repeat homology <AN17>
F:630-662/Domain: ankyrin repeat homology <AN18>
F:663-695/Domain: ankyrin repeat homology <AN19>
F:696-728/Domain: ankyrin repeat homology <AN20>
F:729-761/Domain: ankyrin repeat homology <AN21>
F:762-794/Domain: ankyrin repeat homology <AN22>
F:795-827/Domain: ankyrin repeat homology <AN23>

Query Match      2.2%; Score 12; DB 2; Length 4377;
Best Local Similarity 100.0%; Pred. No. 0.027;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 393 TTTTITTTTTTS 404
    |||||||
```

Db 3968 TTTT TTTT TTTT TTTT TTTT 3979

RESULT 13

alkaline phosphatase (EC 3.1.3.1), intestinal type II - rat (fragment)
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 16-Aug-2004
 C:Accession: B56888
 R:Engle, M.J.; Alpers, D.H.
 Clin. Chem. 38, 2506-2509, 1992
 A:Title: The two mRNAs encoding rat intestinal alkaline phosphatase represent two distinct
 A:Reference number: A56888; MUID:93092310; PMID:1458592
 A:Accession: B56888
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-67 <ENG>
 A:Experimental source: duodenal mucosa
 A:Note: sequence extracted from NCBI backbone (NCBIN:121249, NCBIP:121252)
 C:Superfamily: Alkaline phosphatase
 C:Keywords: intestine; membrane protein; phosphoric monoester hydrolase

Query Match 2.0%; Score 11; DB 2; Length 67;
 Best Local Similarity 100.0%; Pred. No. 0.0065;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 393 TTTT TTTT TTTT TTTT TTTT 403
 |||||
 Db 27 TTTT TTTT TTTT TTTT TTTT 37

RESULT 14

hypothetical protein Y43F8C.9 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 C:Accession: T26880
 R:Ainscough, R.
 submitted to the EMBL Data Library, October 1998
 A:Reference number: Z20279
 A:Accession: T26880
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: DNA
 A:Residues: 1-108 <WIL>
 A:Cross-references: UNIPROT:O9XWN0; EMBL:AL032637; PIDN:CAA21621.1; CESP:Y43F8C.9
 A:Experimental source: Clone Y43F8C
 C:Genetics:
 A:Gene: CESP:Y43F8C.9
 A:Introns: 40/3

Query Match 2.0%; Score 11; DB 2; Length 108;
 Best Local Similarity 100.0%; Pred. No. 0.0098;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 393 TTTT TTTT TTTT TTTT TTTT 403
 |||||
 Db 50 TTTT TTTT TTTT TTTT TTTT 60

RESULT 15

hypothetical protein Y24F12A.d - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 08-Sep-2000
 C:Accession: T26561
 R:Lennard, N.
 submitted to the EMBL Data Library, September 1999
 A:Reference number: Z20233
 A:Accession: T26561
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: DNA
 A:Residues: 1-164 <WIL>
 A:Cross-references: EMBL:AL110480; PIDN:CAB54380.1; CESP:Y24F12A.d

A:Experimental source: clone Y24F12A
 C:Genetics:
 A:Gene: CESP:Y24F12A.d
 A:Introns: 137/1
 C:Superfamily: Caenorhabditis elegans hypothetical protein Y9D1A.2

Query Match 2.0%; Score 11; DB 2; Length 164;
 Best Local Similarity 100.0%; Pred. No. 0.014;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 393 TTTT TTTT TTTT TTTT TTTT 403
 |||||
 Db 112 TTTT TTTT TTTT TTTT TTTT 122

Search completed: May 13, 2005, 12:07:18
 Job time : 48 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 13, 2005, 11:15:16 ; Search time 168 Seconds
(without alignments)
1284.597 Million cell updates/sec

Title: US-09-189-415b-11
Perfect score: 2840
Sequence: 1 MPIGNLGHNPVNNISPPAP.....SNSAVNTNPPAPGSHRFV 558

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq 16Dec04: *
1: Geneseqp1980s: *
2: Geneseqp1990s: *
3: Geneseqp2000s: *
4: Geneseqp2001s: *
5: Geneseqp2002s: *
6: Geneseqp2003as: *
7: Geneseqp2003bs: *
8: Geneseqp2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2840	100.0	558	ADC00799	Enterohae
2	2624.5	92.4	559	2 AAY06221	Aay06221 EHEC E. c
3	1534.5	54.0	549	2 AAY06220	Aay06220 EPEC E. c
4	406	14.3	107	3 AAB20576	Aab20576 Intimin C
5	187	6.6	1300	4 ABB68075	Abb68075 Drosophil
6	182.5	6.4	2586	4 ABB66878	Abb66878 Drosophil
7	180	6.3	596	3 AAY99408	Aay99408 Human PRO
8	180	6.3	596	4 AAB66157	Aab66157 Protein o
9	180	6.3	596	4 AAU29178	Aau29178 Human PRO
10	180	6.3	596	4 AAB87575	Aab87575 Human PRO
11	180	6.3	596	5 ABG95900	Abg95900 Human sec
12	180	6.3	596	6 ABUS8554	Abus8554 Human PRO
13	180	6.3	596	6 ABUS88102	Abus88102 Novel hum
14	180	6.3	596	6 ABUS84417	Abus84417 Human sec
15	180	6.3	596	6 ABR66291	Abr66291 Human sec
16	180	6.3	596	6 ABR65681	Abr65681 Human sec
17	180	6.3	596	6 ABUS9621	Abus9621 Human sec
18	180	6.3	596	6 ABUS2860	Abus2860 Human PRO
19	180	6.3	596	6 ABUS9981	Abus9981 Novel hum
20	180	6.3	596	6 ABR68230	Abr68230 Human sec
21	180	6.3	596	6 ABUS96283	Abus96283 Novel hum
22	180	6.3	596	6 ABUS92714	Abus92714 Human sec
23	180	6.3	596	6 ABO08791	Abo08791 Human sec
24	180	6.3	596	6 ABO02843	Abo02843 Human sec
25	180	6.3	596	6 ABR74997	Abr74997 Human sec

26	180	6.3	596	6 ABR94759	AbR94759 Human PRO
27	180	6.3	596	6 ABUS8732	Abu85732 Human PRO
28	180	6.3	596	6 ABUS8892	Abu8892 Novel hum
29	180	6.3	596	6 ABUS98107	Abu98107 Novel hum
30	180	6.3	596	6 ABUS91813	Abu91813 Novel hum
31	180	6.3	596	6 ABUS9506	Abu89506 Human PRO
32	180	6.3	596	6 ABUS86347	Abu86347 Human PRO
33	180	6.3	596	6 ABUS67560	Abu67560 Human PRO
34	180	6.3	596	6 ABUS0588	Abu0588 Human PRO
35	180	6.3	596	6 ABUS0925	Abu0925 Novel hum
36	180	6.3	596	6 ABO33984	Abo33984 Human sec
37	180	6.3	596	6 ABR99506	AbR99506 Human sec
38	180	6.3	596	6 ABR98896	AbR98896 Human sec
39	180	6.3	596	6 ABO16419	Abo16419 Human sec
40	180	6.3	596	6 ABR92319	AbR92319 Human sec
41	180	6.3	596	6 ABO18960	Abo18960 Human sec
42	180	6.3	596	6 ABR78381	AbR78381 Human sec
43	180	6.3	596	6 ABUS72001	Abu72001 Novel hum
44	180	6.3	596	6 ABUS5117	Abu5117 Novel hum
45	180	6.3	596	6 ABO00256	Abo00256 Novel hum

ALIGNMENTS

RESULT 1

ADC00799
ID ADC00799 standard; protein; 558 AA.

XX

AC ADC00799;

XX

DT 04-DEC-2003 (first entry)

XX

DE Enterohaemorrhagic E. coli O157:H7-specific protein SEQ ID NO: 844.

XX

KW enterohaemorrhagic; anti-bacterial.

XX

OS Escherichia coli; O157:H7.

XX

PN JP2002355074-A.

XX

PD 10-DEC-2002.

XX

PF 24-JAN-2002; 2002JP-00015959.

XX

PR 24-JAN-2001; 2001JP-00112010.

XX

PA (UYTS-) UNIV TSUKUBA.

XX

DR WPI; 2003-451640/43.

XX

PT Enterohaemorrhagic Escherichia coli O157:H7-specific nucleic acid molecule and a polypeptide and its use, a polypeptide, a vector and a host cell.

XX

PS Claim 3; SEQ ID NO 844; 2067pp; Japanese.

XX

CC The invention relates to a novel enterohaemorrhagic Escherichia coli O157:H7-specific nucleic acid molecule. A polynucleotide of the invention has anti-bacterial activity. The polypeptide can be used in detection and/or treatment of O157:H7 infection. The nucleotide sequence of the genome of Enterohaemorrhagic E. coli O157:H7 was determined. The present invention represents an E. coli O157:H7-specific polypeptide of the invention.

XX

SQ Sequence 558 AA;

XX

Query Match 100.0%; Score 2840; DB 7; Length 558;

Best Local Similarity 100.0%; Pred. No. 1.4e-193;

Matches 558; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX

QY 1 MPIGNLGHNPVNNISPPAPLPSQTDGAGGRGQLINSTGPGSRALFTPVNNSMADSGD 60

Db 1 MPIGNLGHNPVNNISPPAPLPSQTDGAGGRGQLINSTGPGSRALFTPVNNSMADSGD 60

QY 61 NRASDVPLGVNPMRLAAASEITLNDGFEVLHDHGPLDLNLRIGSSVFRVETQEDGKHIA 120
 DB |||||
 QY 121 VQORNGVTSVVLSDOEYARLOSIDPEGKDFVFTGGRGGAGHAMVTVASDITEARQIL 180
 DB |||||
 QY 121 VQORNGVTSVVLSDOEYARLOSIDPEGKDFVFTGGRGGAGHAMVTVASDITEARQIL 180
 DB |||||
 QY 181 ELLEPKGTGESKAGESKGVGELRESNCAENTTETQSTSTSSLRSDPKLWALGTAT 240
 DB |||||
 QY 241 GLIGLAATGIVQALALTPEDSPPTTDPDAAASATETATRDQLTKEAFQNPQKNIDE 300
 DB |||||
 QY 301 LGNAIPSGVLKDDVVANIEEQAAGEAKQAENNAQAQKYEQAQKQEEELKVSSG 360
 DB |||||
 QY 301 LGNAIPSGVLKDDVVANIEEQAAGEAKQAENNAQAQKYEQAQKQEEELKVSSG 360
 DB |||||
 QY 361 AGVGLSGALILGGGIGVAVTAALHRKNQVPEOTTTTTTTTTTSARTVENKANNTPAQG 420
 DB |||||
 QY 361 AGVGLSGALILGGGIGVAVTAALHRKNQVPEOTTTTTTTTTTSARTVENKANNTPAQG 420
 DB |||||
 QY 421 NVDTFGSEDTMESRRSSMASTSTFFDTSSIGTVQNPYADVKTSLHDSQVPTSNSNTSVQ 480
 DB |||||
 QY 481 NMGNTDSVYVSTIQHPPRDTTNGARLLGNPSAGIQSTYARLALSGGLRHDMMGGILTGGSN 540
 DB |||||
 QY 541 SAVNTSNPPAPGSHRFV 558
 DB |||||
 QY 541 SAVNTSNPPAPGSHRFV 558
 DB |||||

RESULT 2
 ID AAY06221 standard; protein; 559 AA.
 XX AAY06221;
 AC AAY06221;
 DT 16-AUG-1999 (first entry)
 DE EHEC E. coli translocated intimin receptor (Tir).
 XX Tir; translocated intimin receptor; Hp90; enterohaemorrhagic; EHEC;
 KW infection; diagnosis; vaccine.
 XX Escherichia coli.
 OS
 FH Key Location/Qualifiers
 FT Misc-difference 453
 FT /note= "encoded by codon of 1 apparent nucleotide,
 FT causing frameshift in the DNA sequence"
 XX
 PN WO9224576-A1.
 XX
 XX 20-MAY-1999.
 PD
 XX 10-NOV-1998; 98WO-CA001042.
 PF
 XX 12-NOV-1997; 97US-0065130P.
 PR
 XX (UYBR-) UNIV BRITISH COLUMBIA.
 PA
 XX Finlay BB, Kenny B, Devinney R, Stein M;
 PI
 XX WPI: 1999-337712/28.
 DR N-PSDB; AAX58859.
 XX

PT New translocated intimin receptor useful for treating infection by
 PT enteropathogenic or enterohaemorrhagic Escherichia coli.
 XX
 PS Claim 7; Page 55-58; 91pp; English.
 XX
 CC The present sequence represents Tir, a novel translocated intimin
 CC receptor (formerly termed Hp90) from an enterohaemorrhagic Escherichia
 CC coli (EHEC) strain. The sequence was deduced from an isolated Tir
 CC polynucleotide (see AAX58859). Tir proteins are secreted by attaching and
 CC effacing pathogens such as EHEC and EPEC (see AAY06220) E. coli. The
 CC bacterial pathogens insert their own receptors into mammalian cell
 CC surfaces, to which the pathogen then adheres to trigger additional host
 CC signaling events and actin nucleation. Diagnosis of disease caused by
 CC pathogenic E. coli can be performed by use of antibodies that bind to Tir
 CC to detect the protein or the use of nucleic acid probes for detection of
 CC nucleic acids encoding Tir polypeptide. Isolated Tir nucleic acids, Tir
 CC peptides, a recombinant method for producing recombinant Tir, antibodies
 CC which bind to Tir, and a kit for the detection of Tir-producing E. coli
 CC are provided. A method of immunising a host with Tir to induce a
 CC protective immune response is also provided. In addition, Tir fusion
 CC proteins can be used in attenuated E. coli to induce a cell-mediated
 CC immune response to other polypeptides, e.g. antigens. A method for
 CC screening for compounds which interfere with the binding of bacterial
 CC pathogens to their receptors is further provided
 XX
 SQ Sequence 559 AA;
 Query Match 92.4%; Score 2624.5; DB 2; Length 559;
 Best Local Similarity 94.0%; Pred. No. 3.1e-170;
 Matches 528; Conservative 3; Mismatches 24; Indels 7; Gaps 3;
 QY 1 MPIGNLGHNPVNNISIPAPPPLPSQTDGAGRGQLNSTGPLGSRALTFPVNNSMADSGD 60
 DB |||||
 QY 1 MPIGNLGHNPVNNISIPAPPPLPSQTDGAGRGQLNSTGPLGSRALTFPVNNSMADSGD 60
 DB |||||
 QY 61 NRASDVPLGVNPMRLAAASEITLNDGFEVLHDHGPLDLNLRIGSSVFRVETQEDGKHIA 120
 DB |||||
 QY 61 NRASDVPLGVNPMRLAAASEITLNDGFEVLHDHGPLDLNLRIGSSVFRVETQEDGKHIA 120
 DB |||||
 QY 121 VQORNGVTSVVLSDOEYARLOSIDPEGKDFVFTGGRGGAGHAMVTVASDITEARQIL 180
 DB |||||
 QY 121 VQORNGVTSVVLSDOEYARLOSIDPEGKDFVFTGGRGGAGHAMVTVASDITEARQIL 180
 DB |||||
 QY 181 ELLEPKGTGESKAGESKGVGELRESNCAENTTETQSTSTSSLRSDPKLWALGTAT 240
 DB |||||
 QY 181 ELLEPKGTGESKAGESKGVGELRESNCAENTTETQSTSTSSLRSDPKLWALGTAT 240
 DB |||||
 QY 241 GLIGLAATGIVQALALTPEDSPPTTDPDAAASATETATRDQLTKEAFQNPQKNIDE 300
 DB |||||
 QY 241 GLIGLAATGIVQALALTPEDSPPTTDPDAAASATETATRDQLTKEAFQNPQKNIDE 300
 DB |||||
 QY 301 LGNAIPSGVLKDDVVANIEEQAAGEAKQAENNAQAQKYEQAQKQEEELKVSSG 360
 DB |||||
 QY 301 LGNAIPSGVLKDDVVANIEEQAAGEAKQAENNAQAQKYEQAQKQEEELKVSSG 360
 DB |||||
 QY 361 AGVGLSGALILGGGIGVAVTAALHRKNQVPEOTTTTTTTTTTSARTVENKANNTPAQG 420
 DB |||||
 QY 361 AGVGLSGALILGGGIGVAVTAALHRKNQVPEOTTTTTTTTTTSARTVENKANNTPAQG 420
 DB |||||
 QY 421 NVDTFGSEDTMESRRSSMASTSTFFDTSSIGTVQNPYADVKTSLHDSQVPTSNSNTSV 479
 DB |||||
 QY 421 NVDTFGSEDTMESRRSSMASTSTFFDTSSIGTVQNPYADVKTSLHDSQVPTSNSNTSV 479
 DB |||||
 QY 480 QNM---GNTDSVYVSTIQHPPRDTTNGARLLGNPSAGIQSTYARLALSGGLRHDMMGGIL 536
 DB |||||
 QY 478 RLFRIWGIQISVYVSTIQHPPRDTTNGARLLGNPSAGIQSTYARLALSGGLRHDMMGGIL 537
 DB |||||
 QY 537 GGSNSAVNTSNPPAPGSHRFV 558
 DB |||||
 QY 538 GGSNSAVNTSNPPAPGSHRFV 559
 DB |||||

RESULT 3

Db	1	MPIGNLNNVNGNHLI	PPAPLPSTQDGAARCGTGHLSSTGALGSRSLFSLRNSMADS	60
Qy	59	GNRASDVPLFPVNPRLAA	--SEITLNDGFEVLHDHGLDITLNRQIGSSVFRVETQEDG	116
Db	61	VDSR	--DIPGLTFNPSRLAAATSETCLLGGFEVLHDGGLDITLNTQIGPSAFRVEVQADG	118
Qy	117	KHIAVQORNGVETSVLS	SDQBYARLQSIDPBGKKKFTVTTGGRGAGHAMVTVASDITTEAR	176
Db	119	THAAIIGKNGLEVS	VTLSPQEWSSLSQIDTEGKRFVTTGGRGSGHPMTVTVASDIAEAR	178
Qy	177	ORILELLPKPGTG	-----ESKGAGESKVGCELRESNSGAENTTETOTSTSTSLRS	227
Db	179	TRILAKLPDHNHGG	RQPKDVTFRSVGVGSAGI-----DDGV--VSEHTITNSSVRS	230
Qy	228	DPKLMALGTVATGLIGLAAT	GIVOALALTEPPDSPTTTTDPDAAASATETATRDQLTKEA	287
Db	231	DPKFWVS	GAIAAGLAGLAATGIAQALALTEPPDDPTTTDPDQAAANAESATKDKLTQEA	290
Qy	288	FONPDNOKWNIDELGNA	IPSGVLKDDVVANTEEQAAGBEAKQOATENNAQAOKKYDEQ	347
Db	291	FXNPENQKVNIDANGNA	IPSGELKDDIVEQIAQAQKEAGEVARQQAQVESAQAQRYEDQ	350
Qy	348	QAKROEELKVSSGAGYGLSGAL	ILGGGIGVAVTAALHRKNQPVETTTTTTTTTTTTSART	407
Db	351	HARROEELQLSSGIGYGLSSALI	VAGGIGAGVTTLALHRRNPQAEQTITTTT-----HT	403
Qy	408	VENKPANNTPAQGNVDTF	GESDFTMESRRSSMASTSTFTDTSSTIGTVQNPYADV-----K	462
Db	404	VVQQTGGIPQHKVALM	PEQERRFRDRDSQGSVASTHWSDS--SEVVNFAEVGGARNS	462
Qy	463	TSLHD	-----SQVPTSNSTSVOMGNNTDSVYVSTTQHPPRDITDNGARLLGNPSAGIQ	516
Db	463	LSAQHPEEHYDEVAAD	PGYSVIQNFSGSGPVT-----GRLLGTGGGQIQ	507
Qy	517	STYARLALSGLRHDMG	LTCGNSAVNTSNPPAPGSHRFV	558
Db	508	STYALLANSGLRLGMGL	TSGGTAVSSVNAATPGPVRFV	549
RESULT 4				
ID	AAB20576	standard; protein; 107 AA.		
XX	AC	AAB20576;		
XX	DT	08-DEC-2000	(first entry)	
XX	DE	Intimin C-terminal Tir binding domain amino acid sequence.		
XX	KW	Intimin; Tir binding domain; inhibitor; intimin adhesion; screening;		
XX	KW	Tir-independent eukaryotic cell binding activity; bacterial infection;		
XX	OS	diarrhoea; antibacterial.		
XX	PN	Unidentified.		
XX	PD	WO200045173-A1.		
XX	PP	03-AUG-2000.		
XX	PP	31-JAN-2000; 2000WO-GB000254.		
XX	PR	29-JAN-1999; 99GB-00001897.		
XX	PA	(IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.		
XX	PI	Frankel GM, Matthews SJ, Hale CB, Dougan G;		
XX	DR	WPI; 2000-499357/44.		
XX	PT	Screening for inhibitors of intimin binding to eukaryotic cells, for use		
XX	PT	in diagnosing, preventing and treating bacterial infections, especially		
XX	PT	Escherichia coli O157:H7.		

AA06220	1	MEIGNLNNVNGNHLI	PPAPLPSTQDGAARCGTGHLSSTGALGSRSLFSLRNSMADS	58
XX	ID	AA06220	standard; protein; 549 AA.	
XX	AC	AA06220;		
XX	DT	16-AUG-1999	(first entry)	
XX	DE	EEPC E. coli	translocated intimin receptor (Tir).	
XX	DE	Tir;	translocated intimin receptor; Hp90; enteropathogenic; EPEC;	
KW	KW	infection; diagnosis;	vaccine.	
XX	OS	Escherichia coli.		
XX	EH	Key	Location/Qualifiers	
FT	FT	Misc-difference 180	/note= "encoded by AAA"	
FT	FT	Domain	234. .253	
FT	FT	Misc-difference 314	/note= "putative transmembrane domain"	
FT	FT	Domain	364. .386	
FT	FT	Domain	/note= "putative transmembrane domain"	
XX	PN	WO9924576-A1.		
XX	PD	20-MAY-1999.		
XX	XX	10-NOV-1998;	98WO-CA001042.	
XX	PF	12-NOV-1997;	97US-0065130P.	
PR	PR	(UYBR-) UNIV BRITISH COLUMBIA.		
XX	PI	Finlay BB, Kenny B, Devinney R, Stein M;		
XX	XX	WPI: 1999-337712/28.		
XX	DR	N-PSDB; AA058858.		
XX	PT	New translocated intimin receptor useful for treating infection by enteropathogenic or enterohemorrhagic Escherichia coli.		
XX	PS	Claim 6; Page 55-58; 91pp; English.		
XX	CC	The present sequence represents Tir, a novel translocated intimin receptor (formerly termed Hp90) from an enteropathogenic Escherichia coli (EPEC) strain. The sequence was deduced from an isolated tir polynucleotide (see AA058858). Tir proteins are secreted by attaching and effacing pathogens such as EPEC and EHEC (see AA06221) E. coli. The bacterial pathogens insert their own receptors into mammalian cell surfaces, to which the pathogen then adheres to trigger additional host signaling events and actin nucleation. Diagnosis of disease caused by pathogenic E. coli can be performed by use of antibodies that bind to Tir to detect the protein or the use of nucleic acid probes for detection of nucleic acids encoding Tir polypeptide. Isolated Tir nucleic acids, Tir peptides, a recombinant method for producing recombinant Tir, antibodies which bind to Tir, and a kit for the detection of Tir-producing E. coli are provided. A method of immunising a host with Tir to induce a protective immune response is also provided. In addition, Tir fusion proteins can be used in attenuated E. coli to induce a cell-mediated immune response to other polypeptides, e.g. antigens. A method for screening for compounds which interfere with the binding of bacterial pathogens to their receptors is further provided		
XX	XX	Sequence 549 AA;		
XX	XX	Query Match	54.0%; Score 1534.5; DB 2; Length 549;	
XX	XX	Best Local Similarity	56.4%; Pred. No. 1.3e-100;	
XX	XX	Matches 328; Conservative	60; Mismatches 137; Indels 57; Gaps 11;	
Qy	1	MEIGNLNNVNGNHLI	PPAPLPSTQDGAARCGTGHLSSTGALGSRSLFSLRNSMADS	58


```

QY 525 SGGLRDMGGLTGGNSAVNTSNNPPAPG 553
Db 475 AGSGTAALTGMMHTTSHA-STAVSEAKPG 502

RESULT 8
AAB66157
ID AAB66157 standard; protein; 596 AA.
XX
AC AAB66157;
XX
DT 02-APR-2001 (first entry)
DE Protein of the invention #69.
XX
KW Secreted; transmembrane; gene therapy.
XX
OS Unidentified.
XX
PN WO200078961-A1.
XX
PD 28-DEC-2000.
XX
PF 18-FEB-2000; 2000WO-US004342.
XX
PR 23-JUN-1999; 99US-0141037P.
PR 20-JUL-1999; 99US-0144758P.
PR 26-JUL-1999; 99US-0145698P.
PR 01-SEP-1999; 99WO-US020111.
PR 29-OCT-1999; 99US-0162506P.
PR 30-NOV-1999; 99WO-US028313.
PR 02-DEC-1999; 99WO-US028551.
PR 16-DEC-1999; 99WO-US030095.
PR 05-JAN-2000; 2000WO-US000219.
PR 06-JAN-2000; 2000WO-US000376.
XX
PA (GETH ) GENENTECH INC.
XX
PI Baker KP, Botstein D, Desnoyers L, Eaton DL, Ferrara N, Fong S;
PI Gao W, Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ;
PI Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tumas D, Watanabe CK;
PI Williams PM, Wood WI;
XX
DR WPI; 2001-071395/08.
XX
PT Secreted and transmembrane proteins and nucleic acids designated PRO,
PT useful as hybridization probes, in chromosome and gene mapping and gene
PT therapy.
XX
PS Claim 1; Fig 138; 787pp; English.
XX
CC The present invention relates to secreted and transmembrane proteins.
CC These proteins and the DNA encoding them may be used as hybridization
CC probes, in chromosome and gene mapping and in the generation of anti-
CC sense RNA and DNA. They may also be used to generate either
CC transgenic animals or knockout animals which are in turn useful for
CC development and screening of therapeutically useful reagents. The nucleic
CC acids may also be used in gene therapy
XX
SQ Sequence 596 AA;

Query Match 6.3%; Score 180; DB 4; Length 596;
Best Local Similarity 19.2%; Pred. No. 0.00044;
Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;

QY 4 GNIGHNPVNNSTPPAPPSPQTDGAGRGQLINSTGPLGSRALFFPVNSWADSGDRA 63
Db 13 GLLHLEAATNS-----NETSANTGSSVISSG-----ASTATNSG 49
QY 64 SDYPGLVPVPMRLAASEITINDGFEVLHDHGLDITLNRQIGSSV-FRVETQEDGKHIAVG 122
Db 50 SSVTSSGVSATISGSSVTSN-GVSIV-TNSEPHITSSGISTATNSEFSTAGSGISIATN 107

```

```

QY 123 QRNGVETS--VVLSDQEVARLOSIDPEGKDKVFVFTGRCGAGHAMVTVASDITEARQRL 180
Db 108 SESSTTSSGASTATNSE-----SSTPSSGASTVTNCGSSVTSSGASTATNSESTVSSRAS 163
QY 181 ELLEPKGTGESKGAGSKGVGELRESNCGAENVTETOTSTSTSSLSRSDPKLMLALGTVA 240
Db 164 TATNSESSTLSSGASTATN-SDSSTTSSGASTATNSESTTSSG-----ASTAT 211
QY 241 GLIGLAATGIVQALALTPEDSPFTTDPDAASATETATRDQLTKEAFQNPQNKYNIDE 300
Db 212 N-----SESSTVSSRASTATNSESTT---SSGASTATNSESTTSSGAGTATNSESTTS 264
QY 301 LG-----NAIPSGVLKDDVAVNIEBOAKAAGEAKQAENNAQAKKYDEQOAKR 351
Db 265 SGASTATNSDSSTVSSGA---STATNSESTTSSGAST-----ATN 302
QY 352 QEELKVSSGAGYGL-SGALLILGGIGVAVTAALHRKNQPVQETTTTTTTTTTTSARTVEN 410
Db 303 SESSTTSSGASTATNSDSSTTSSGAGTATNSESTTSSGASTATNSESTTSSGASTATN 362
QY 411 KPANNTPAGNVDTFGESEDTMESRRSSMASTSTFTDTSISICTVQNPYADVKTSS----- 464
Db 363 SESSTTSSGANTAT-NSESSTVSSGASTATNSESTTSSGASTATNSESTTSSGASTAT 421
QY 465 LHDSQVPTNSNTSVQNMGNNTSVVYTIQHPPTDTTNGARLLGNPSAGIOSTYARLAL 524
Db 422 NSDSSTTSSGASTATNSESTTSSGASTATNSESTTSSGASTATNSESTTSSGASTAT 474
QY 525 SGGLRDMGGLTGGNSAVNTSNNPPAPG 553
Db 475 AGSGTAALTGMMHTTSHA-STAVSEAKPG 502

RESULT 9
AAU29178
ID AAU29178 standard; protein; 596 AA.
XX
AC AAU29178;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human PRO polypeptide sequence #155.
XX
KW PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep;
KW dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;
KW blood; chondrocyte cell; cell proliferation; cell differentiation; colon;
KW adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200168848-A2.
XX
PD 20-SEP-2001.
XX
PF 28-FEB-2001; 2001WO-US006520.
XX
PR 01-MAR-2000; 2000WO-US005601.
PR 02-MAR-2000; 2000WO-US005841.
PR 03-MAR-2000; 2000US-0187202P.
PR 06-MAR-2000; 2000US-0186968P.
PR 14-MAR-2000; 2000US-0189320P.
PR 14-MAR-2000; 2000US-0189328P.
PR 15-MAR-2000; 2000WO-US006884.
PR 21-MAR-2000; 2000US-0190828P.
PR 21-MAR-2000; 2000US-0191007P.
PR 21-MAR-2000; 2000US-0191048P.
PR 21-MAR-2000; 2000US-0191314P.
PR 28-MAR-2000; 2000US-0192655P.
PR 29-MAR-2000; 2000US-0193032P.
PR 29-MAR-2000; 2000US-0193053P.
PR 30-MAR-2000; 2000WO-US008439.
PR 04-APR-2000; 2000US-0194449P.

```


30-MAR-2000; 2000WO-US008439.
 22-MAY-2000; 2000WO-US014042.
 02-JUN-2000; 2000WO-US015264.
 23-AUG-2000; 2000WO-US023522.
 24-AUG-2000; 2000WO-US023328.
 10-NOV-2000; 2000WO-US030873.
 01-DEC-2000; 2000WO-US032378.
 20-DEC-2000; 2000WO-US034956.
 28-FEB-2001; 2001WO-US006520.
 01-MAR-2001; 2001WO-US006666.
 30-MAY-2001; 2001WO-US017443.
 01-JUN-2001; 2001WO-US017800.
 20-JUN-2001; 2001WO-US019692.
 29-JUN-2001; 2001WO-US021066.
 09-JUL-2001; 2001WO-US021735.
 (GETH) GENENTECH INC.
 Eaton DL, Filvaroff E, Gerritsen ME, Goddard A, Godowski PJ;
 Grimaldi JC, Gurney AL, Watanabe CK, Wood WI;
 N-PSDB; ABS74427.
 WPI: 2002-731348/79.
 DR N-PSDB; ABS74427.
 XX
 PT New isolated secreted and transmembrane PRO polypeptide useful for
 PT modulating biological activity of a cell, or for treating sports-related
 PT joint problems, osteoarthritis or rheumatoid arthritis.
 XX
 PS Claim 20; Fig 100; 399pp; English.
 XX
 CC The invention relates to an isolated secreted and transmembrane PRO
 CC polypeptide having 80 % sequence identity to a sequence appearing as
 CC ABG95851-ABG95934 or their associated signal peptide, or a sequence of an
 CC extracellular domain of the proteins with their associated signal peptide
 CC or lacking its associated signal peptide. Also included are the nucleic
 CC acids encoding the proteins, vectors, host cells, fusion proteins and
 CC antibodies which specifically bind to the proteins. The proteins are
 CC useful for detecting a polypeptide designated as A, B, C or D in a sample
 CC suspected of containing A, B, C or D polypeptide, by contacting the
 CC sample with a polypeptide designated as E, F, G, H or I (or vice versa)
 CC and determining the formation of a A/E, B/F, B/G, C/H or D/I polypeptide
 CC conjugate in the sample, where the formation of the conjugate is
 CC indicative of the presence of an A, B, C or D polypeptide in the sample,
 CC where A is a PRO10272 polypeptide, B is a PRO20110 polypeptide, C is a
 CC PRO10096 polypeptide, D is a PRO19760 polypeptide, E is a PRO5801
 CC polypeptide, F is a PRO1 polypeptide, G is a PRO20040 polypeptide, H is a
 CC PRO20233 polypeptide and I is a PRO1890 polypeptide. The sample comprises
 CC a cell suspected of expressing the A, B, C or D polypeptide. The E, F, G,
 CC H or I polypeptide is labeled with a detectable label or is attached to a
 CC solid support. The proteins are useful for linking a bioactive molecule
 CC to a cell expressing a polypeptide designated as A, B, C or D or E, F, G,
 CC H or I. The bioactive molecule is a toxin, a radiolabel or an antibody.
 CC The bioactive molecule causes death of the cell. A, B, C, D, E, F, G, H,
 CC or I, or antibodies against them are useful for modulating a biological
 CC activity of a cell expressing a polypeptide designated as A, B, C or D or
 CC E, F, G, H, or I. The cell is killed. The proteins are useful for
 CC identifying agonists or antagonists, for the preparation of a medicament
 CC useful in the treatment of a condition, for the preparation of a medicament
 CC proteins, as molecular weight markers for protein electrophoresis
 CC purposes, and as therapeutic agents for treating sports-related joint
 CC problems, articular cartilage defects, osteoarthritis or rheumatoid
 CC arthritis. Nucleic acids encoding the proteins are useful as
 CC hybridisation probes, in chromosome and gene mapping, in the generation
 CC of anti-sense RNA and DNA, for the preparation of the proteins, to
 CC generate transgenic or knockout animals which are useful in the
 CC development and screening of therapeutic useful reagents, for chromosome
 CC identification, and in gene therapy. The antibody is useful as a
 CC therapeutic agent, in a diagnostic assay and for affinity purification of
 CC the protein from recombinant cell culture natural sources. The present
 CC sequence represents a novel secreted or transmembrane protein of the
 CC invention
 XX
 SQ Sequence 596 AA;

Query Match
 Best Local Similarity 6.3%; Score 180; DB 5; Length 596;
 Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;
 QY 4 QNLGHNPVNNNSIPAPPPLPSCDTGAGRGQLINSTGPLGRALPTFVRNSWADSGDNRA 63
 DB 13 GLLHLEAATNS-----NETSTANTGSSVITSSG-----ASTATNSG 49
 QY 64 SDVPCLPYNMRLAASEITLNDGEVLHDHPLDPLNFIQIGSSV-FRVETQEDGKHIAVG 122
 DB 50 SSVTSSGVSTATIGSSVTSN-GVSIV-TNSEFHTTSSGISTATNSESTSSGSIATN 107
 QY 123 QRNGVETS--VVLSDQEARLQSIDPEGKDKFVFTGGRGAGHAMVTVASDITEARQRL 180
 DB 108 SESSTSSGASTATNSE-----SSTPSSGASTVTSNGSSVTSNGASTATNSESTSSRAS 163
 QY 181 ELLPEKGTGESKAGESKVGELRESNGAENTTETQTSTSSLRSDPKLWLALGTVAT 240
 DB 164 TATNSESTLSSGASTATN-SDSSTSSGASTATNSESTSSG-----ASTAT 211
 QY 241 GLIGLAATGIQALALTPPEPSPITTDPAASAATETATROLTKAEFQNPDNQKVNIDE 300
 DB 212 N-----SESTVSSRASTATNSESTT---SSGASTATNSESTTSSNGAGTATNSESTTS 264
 QY 301 LG-----NAIPSGVLKDDVVANIEQAKAAGEEAKQAIENNAQAKKYDEQAKR 351
 DB 265 SGASTATNDSSTVSSGA---STATNSESTSSGAST-----ATN 302
 QY 352 QEELKVSAGGYGL-SGALLGGGIGVAVTAALHRKNQFVEQTITTTTTTTSARTVEN 410
 DB 303 SESSTSSGASTATNDSSTSSGASTATNSESTVSSGISTVTSNESSTSSGANTATN 362
 QY 411 KPANNTPAQGVNTPGSEDTWESRRSSMASTSSFTDTSIGTVQNPYADVKT----- 464
 DB 363 SESSTSSGANTAT-NSESSTVSSGASTATNSESTSSGASTATNSESTSSGASTAT 421
 QY 465 LHDQVPTNSNTSVQNNMGNTDSVVYSTIOHPRPDITDNGARLLGNPSAGIQSTVARLAL 524
 DB 422 NSDSTSSSEASTATNSESTVSSGISTVTSNESSTSSGANTATNDSGSSVTS----- 474
 QY 525 SGGLRHDMMGGLTGGNSGAVTNNPPAPG 553
 DB 475 AGSGTAALTGMHTTSHSA-STAVSEAKPG 502

RESULT 12
 ABUS8554
 ID ABUS8554 standard; protein; 596 AA.
 XX AC ABUS8554;
 XX DT 15-APR-2003 (first entry)
 XX DE Human PRO polypeptide #155.
 XX KW Human; PRO; cytostatic; tumour; cancer; breast; lung; stomach; liver;
 KW dog; cat; cow; horse; sheep; pig; goat; rabbit; ADEPT;
 KW antibody-dependent enzyme mediated prodrug therapy.
 XX OS Homo sapiens.
 XX PN US2003027272-A1.
 XX PD 06-FEB-2003.
 XX PF 21-JUN-2002; 2002US-00176492.
 XX PR 18-SEP-1997; 97US-0059263P.
 PR 18-SEP-1997; 97US-0059266P.
 PR 17-OCT-1997; 97US-0062250P.
 PR 21-OCT-1997; 97US-0063486P.
 PR 24-OCT-1997; 97US-0063120P.

PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069333P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 97US-0077450P.
PR 11-MAR-1998; 98US-0077633P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 15-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 97US-0063121P.
PR 11-JUN-1998; 97US-0063540P.
PR 11-JUN-1998; 97US-0063541P.
PR 11-JUN-1998; 97US-0063544P.
PR 12-JUN-1998; 97US-0063564P.
PR 12-JUN-1998; 97US-0063734P.
PR 16-JUN-1998; 97US-0063870P.
PR 16-JUN-1998; 97US-0064103P.
PR 17-JUN-1998; 97US-0065311P.
PR 17-JUN-1998; 97US-0066120P.
PR 17-JUN-1998; 97US-0066466P.
PR 17-JUN-1998; 97US-0066772P.
PR 19-JUN-1998; 97US-0069333P.
PR 19-JUN-1998; 97US-0069425P.
PR 22-JUN-1998; 97US-0069870P.
PR 22-JUN-1998; 97US-0068017P.
PR 24-JUN-1998; 97US-0077450P.
PR 24-JUN-1998; 98US-0077633P.
PR 24-JUN-1998; 98US-0077649P.
PR 24-JUN-1998; 98US-0078886P.
PR 24-JUN-1998; 98US-0078939P.
PR 24-JUN-1998; 98US-0079664P.
PR 25-JUN-1998; 98US-0079786P.
PR 25-JUN-1998; 98US-0080107P.
PR 25-JUN-1998; 98US-0080194P.
PR 25-JUN-1998; 98US-0080327P.
PR 25-JUN-1998; 98US-0080333P.
PR 25-JUN-1998; 98US-0081049P.
PR 25-JUN-1998; 98US-0081070P.
PR 26-JUN-1998; 98US-0081195P.
PR 26-JUN-1998; 98US-0081838P.
PR 26-JUN-1998; 98US-0082568P.
PR 26-JUN-1998; 98US-0082569P.
PR 01-JUL-1998; 98US-0082704P.
PR 02-JUL-1998; 98US-0082797P.
PR 02-JUL-1998; 98US-0083322P.
PR 02-JUL-1998; 98US-0083495P.
PR 02-JUL-1998; 98US-0083496P.
PR 02-JUL-1998; 98US-0083499P.
PR 02-JUL-1998; 98US-0083559P.
PR 04-AUG-1998; 98US-0084366P.
PR 04-AUG-1998; 98US-0084414P.
PR 10-AUG-1998; 98US-0084639P.
PR 10-AUG-1998; 98US-0084640P.
PR 17-AUG-1998; 98US-0084643P.
PR 17-AUG-1998; 98US-0085579P.
PR 17-AUG-1998; 98US-0085580P.
PR 17-AUG-1998; 98US-0085582P.
PR 17-AUG-1998; 98US-0085700P.
PR 18-AUG-1998; 98US-0086012P.
PR 18-AUG-1998; 98US-0086392P.
PR 18-AUG-1998; 98US-0086486P.
PR 26-AUG-1998; 98US-0087098P.
PR 26-AUG-1998; 98US-0087208P.
PR 26-AUG-1998; 98US-0087609P.
PR 26-AUG-1998; 98US-0087759P.
PR 26-AUG-1998; 98US-0087827P.
PR 26-AUG-1998; 98US-0088025P.
PR 26-AUG-1998; 98US-0088028P.
PR 01-SEP-1998; 98US-0088029P.
PR 01-SEP-1998; 98US-0088033P.
PR 02-SEP-1998; 98US-0088326P.
PR 02-SEP-1998; 98US-0088167P.
PR 02-SEP-1998; 98US-0088202P.
PR 09-SEP-1998; 98US-0088212P.
PR 10-SEP-1998; 98US-0088217P.
PR 10-SEP-1998; 98US-0088655P.
PR 10-SEP-1998; 98US-0088722P.
PR 10-SEP-1998; 98US-0088738P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088866P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 02-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 24-JUL-1998; 98US-0091632P.
PR 04-AUG-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 01-SEP-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.

```
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.
PR 07-OCT-1998; 98US-0016897B.

Query Match
Beat Local Similarity 6.3%; Score 180; DB 6; Length 596;
Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;

QY 4 GNLGNPNVNNISIPAPLPSCDTCAGCGQLINSGTGLGREALTPVRNMAUDSGDNRA 63
DB 13 GULLHLEAATNS-----NETSANTGSSVSSG-----ASTATNSG 49

QY 64 SDVPLGNPMRLAASEITLNDGFVLHDHGLDPLNRIQISSV-FRVETQEDGKHIAVG 122
DB 50 SSVTSSGVSTATISGSSVTSN-GVSIV-TNEFHHTTSGISTATNSESTASSGISIATN 107

QY 123 QRNGVETS--VVLSDQEVARIQSIDPEKDKFVFTGGRGGAGHAMVTVASDITEARQIRL 180
DB 108 SESSTTSSGASTATNSE----SSTPSSGASTVNSGSSVTSSGASTATNSESTVSSRAS 163

QY 181 ELLEPKGTGESKGACESKGVGELRESNGAENTTETQSTSSLSRSDPKLWLALGTVAT 240
DB 164 TATNSESTLSSGASTATN-SDSSTTSSGASTATNSESTTSSG-----ASTAT 211

QY 241 GLIGLAATGIVQALALTPEPDSPPTTDPDAAASATETATRDQLTKEAFQNPDKVNI 300
DB 212 N-----SESTVSSRATATNSESTT---SSGASTATNSESTTSSGAGATATNSESTTS 264

QY 301 LG-----NAIPSGVLKDDVANIIEQAKAAGEAKQAQIENNAQAQKYEQQAKR 351
DB 265 SGASTATNSDSTVSSGA---STATNSESTTSSGAST-----ATN 302

QY 352 QEELKVSGAGVGL-SGALLGGGIGVAVTAALHRKNPVQQTITTTTTTTTSARTVEN 410
DB 303 SESSTTSSGASTATNSDSTSSGAGTATNSESTVSSGISTVNTNSESTPSSGANTATN 362

QY 411 KPAANTPAQGNVDTPGSEDPTMESRRSSMASSTPFDTSSTGTVQNPYADVKTSS----- 464
DB 363 SESSTTSSGANTAT-NSSESTVSSGASTATNSESTTSSGASTATNSESTTSSGASTAT 421

QY 465 LHDSQVPTNSNGTSVONMGNTDSVVYSTIQHPPRDTTNGARLLGNPSAGIQSTYARLAL 524
DB 422 NSDSSTTSSSEASTATNSESTVSSGISTVNTNSESTTSSGANTATNSGSSVTS----- 474

QY 525 SGGLRHDHMGGLTGGNSAVNTSNPPAQ 553
DB 475 AGSGTAALTGMHTTSHSA-STAVSEAKFG 502
```

```
RESULT 13
ABU88102
ID ABU88102 standard; protein; 596 AA.
XX
AC ABU88102;
XX
DT 07-JUL-2003 (first entry)
XX
DE Novel human secreted and transmembrane protein PRO1342.
XX
KW Human; secreted and transmembrane protein: PRO; gene therapy;
KW tumour necrosis factor-alpha release; TNF-alpha release;
KW chondrocyte proliferation; chondrocyte differentiation; tumour;
KW adrenal tumour; lung tumour; colon tumour; breast tumour;
KW prostate tumour; rectal tumour; cervical tumour; liver tumour.
XX
OS Homo sapiens.
XX
PN US2003032127-A1.
XX
PD 13-FEB-2003.
XX
PF 26-JUN-2002; 2002US-00183012.
XX
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0082250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0086120P.
PR 24-NOV-1997; 97US-0066466P.
PR 11-DEC-1997; 97US-0066772P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 08-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 15-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 21-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
```


PR	07-MAY-1998;	98US-0084639P.	PR	10-AUG-1998;	98US-0095998P.
PR	07-MAY-1998;	98US-0084640P.	PR	10-AUG-1998;	98US-0096012P.
PR	07-MAY-1998;	98US-0084643P.	PR	17-AUG-1998;	98US-0096757P.
PR	15-MAY-1998;	98US-0085579P.	PR	17-AUG-1998;	98US-0096766P.
PR	15-MAY-1998;	98US-0085580P.	PR	17-AUG-1998;	98US-0096867P.
PR	15-MAY-1998;	98US-0085582P.	PR	17-AUG-1998;	98US-0096891P.
PR	15-MAY-1998;	98US-0085700P.	PR	17-AUG-1998;	98US-0096897P.
PR	18-MAY-1998;	98US-0086023P.	PR	18-AUG-1998;	98US-0096949P.
PR	22-MAY-1998;	98US-0086392P.	PR	18-AUG-1998;	98US-0096959P.
PR	22-MAY-1998;	98US-0086486P.	PR	18-AUG-1998;	98US-0097022P.
PR	28-MAY-1998;	98US-0087098P.	PR	26-AUG-1998;	98US-0097552P.
PR	28-MAY-1998;	98US-0087208P.	PR	26-AUG-1998;	98US-0097554P.
PR	02-JUN-1998;	98US-0087609P.	PR	26-AUG-1998;	98US-0097955P.
PR	02-JUN-1998;	98US-0087759P.	PR	26-AUG-1998;	98US-0097971P.
PR	03-JUN-1998;	98US-0087827P.	PR	26-AUG-1998;	98US-0097974P.
PR	04-JUN-1998;	98US-0088025P.	PR	26-AUG-1998;	98US-0098014P.
PR	04-JUN-1998;	98US-0088028P.	PR	01-SEP-1998;	98US-0098716P.
PR	04-JUN-1998;	98US-0088029P.	PR	01-SEP-1998;	98US-0098723P.
PR	04-JUN-1998;	98US-0088033P.	PR	02-SEP-1998;	98US-0098803P.
PR	04-JUN-1998;	98US-0088326P.	PR	02-SEP-1998;	98US-0098821P.
PR	05-JUN-1998;	98US-0088167P.	PR	02-SEP-1998;	98US-0098843P.
PR	05-JUN-1998;	98US-0088202P.	PR	09-SEP-1998;	98US-0099602P.
PR	05-JUN-1998;	98US-0088212P.	PR	10-SEP-1998;	98US-0099741P.
PR	05-JUN-1998;	98US-0088217P.	PR	10-SEP-1998;	98US-0099754P.
PR	09-JUN-1998;	98US-0088655P.	PR	10-SEP-1998;	98US-0099763P.
PR	10-JUN-1998;	98US-0088722P.	PR	10-SEP-1998;	98US-0099812P.
PR	10-JUN-1998;	98US-0088738P.	PR	15-SEP-1998;	98US-0100388P.
PR	10-JUN-1998;	98US-0088740P.	PR	16-SEP-1998;	98US-0100662P.
PR	10-JUN-1998;	98US-0088811P.	PR	16-SEP-1998;	98US-0100664P.
PR	10-JUN-1998;	98US-0088824P.	PR	16-SEP-1998;	98US-0101751P.
PR	10-JUN-1998;	98US-0088825P.	PR	16-SEP-1998;	98US-0101751P.
PR	10-JUN-1998;	98US-0088826P.	PR	17-SEP-1998;	98US-0100683P.
PR	11-JUN-1998;	98US-0088861P.	PR	17-SEP-1998;	98US-0100684P.
PR	11-JUN-1998;	98US-0088863P.	PR	17-SEP-1998;	98US-0100919P.
PR	11-JUN-1998;	98US-0088876P.	PR	17-SEP-1998;	98US-0100930P.
PR	12-JUN-1998;	98US-0089090P.	PR	18-SEP-1998;	98US-0100849P.
PR	12-JUN-1998;	98US-0089105P.	PR	18-SEP-1998;	98US-0101014P.
PR	16-JUN-1998;	98US-0089512P.	PR	18-SEP-1998;	98US-0101068P.
PR	16-JUN-1998;	98US-0089514P.	PR	23-SEP-1998;	98US-0101471P.
PR	17-JUN-1998;	98US-0089538P.	PR	23-SEP-1998;	98US-0101472P.
PR	17-JUN-1998;	98US-0089598P.	PR	23-SEP-1998;	98US-0101475P.
PR	17-JUN-1998;	98US-0089653P.	PR	23-SEP-1998;	98US-0101477P.
PR	18-JUN-1998;	98US-0089908P.	PR	24-SEP-1998;	98US-0101739P.
PR	19-JUN-1998;	98US-0089952P.	PR	24-SEP-1998;	98US-0101739P.
PR	22-JUN-1998;	98US-0090246P.	PR	24-SEP-1998;	98US-0101743P.
PR	22-JUN-1998;	98US-0090252P.	PR	25-SEP-1998;	98US-0101786P.
PR	22-JUN-1998;	98US-0090254P.	PR	25-SEP-1998;	98US-0102207P.
PR	24-JUN-1998;	98US-0090429P.	PR	29-SEP-1998;	98US-0102240P.
PR	24-JUN-1998;	98US-0090435P.	PR	29-SEP-1998;	98US-0102330P.
PR	24-JUN-1998;	98US-0090444P.	PR	29-SEP-1998;	98US-0102331P.
PR	24-JUN-1998;	98US-0090461P.	PR	30-SEP-1998;	98US-0102487P.
PR	24-JUN-1998;	98US-0090535P.	PR	30-SEP-1998;	98US-0102570P.
PR	25-JUN-1998;	98US-0090540P.	PR	30-SEP-1998;	98US-0102571P.
PR	25-JUN-1998;	98US-0090676P.	PR	01-OCT-1998;	98US-0102684P.
PR	25-JUN-1998;	98US-0090678P.	PR	01-OCT-1998;	98US-0102687P.
PR	25-JUN-1998;	98US-0090688P.	PR	02-OCT-1998;	98US-0102965P.
PR	25-JUN-1998;	98US-0090690P.	PR	06-OCT-1998;	98US-0103258P.
PR	25-JUN-1998;	98US-0090694P.			
PR	25-JUN-1998;	98US-0090695P.			
PR	25-JUN-1998;	98US-0090696P.			
PR	26-JUN-1998;	98US-0090699P.			
PR	26-JUN-1998;	98US-0090862P.			
PR	26-JUN-1998;	98US-0090863P.			
PR	26-JUN-1998;	98US-0091010P.			
PR	01-JUL-1998;	98US-0091359P.			
PR	01-JUL-1998;	98US-0091544P.			
PR	02-JUL-1998;	98US-0091478P.			
PR	02-JUL-1998;	98US-0091486P.			
PR	02-JUL-1998;	98US-0091626P.			
PR	02-JUL-1998;	98US-0091628P.			
PR	02-JUL-1998;	98US-0091632P.			
PR	24-JUL-1998;	98US-0094006P.			
PR	04-AUG-1998;	98US-0095282P.			

Query Match 6.3%; Score 180; DB 6; Length 596;

Best Local Similarity 19.2%; Pred. No. 0.00044;

Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;

QY 4 GNLGHPNPNVNSIPPAPLPSTODGAGRGQLINSTGPLGRALFTPVNSMADSGDNRA 63

Db 13 GLLHLEAATNS-----NETSTANTGSSVSSG-----ASTATNSG 49

QY 64 SDVPGLPNPNMRLAASEITLNDGFVLHDHGLDITLNRIGSSV-FRVTQEDKGHIAGV 122

Db 50 SSVTSSGVSTATISGSSVTSN-GVSIV-TNSEFHTTSSGISTATNSEFSTASSGISATN 107

QY 123 QNNGVETS--VVLSDQEVARLOSIDPEGKDKVFYTGRCGAGHAMVTVASDITEARQRL 180

Db 108 SESSTTSGASTATNSE-----SSTPSSGASTVTNCGSSVTSSGASTATNSESSTVSSRAS 163
QY 181 ELLEPKGTGSKGAGESKGVGLRBSNSGAENTTETOTSTSTSSLRSDPKLWALGTAT 240
Db 164 TATNSESSTLSSGASTATN-SDSSTTSSGASTATNSESSTSSG-----ASTAT 211
QY 241 GLIGUAATGIVOALALTPEPDSPTTTPDAAAASATETATRDQLTKEAFQNPQNKVNIDE 300
Db 212 N-----SESSTVSSRASTATNSESSTT---SSGASTATNSESSTTSGAGTATNSESSTTS 264
QY 301 LG-----NAIPSGVLKDDVVANIIEQAKAAGEBAKQAIENNAQAQKYDEQQAKR 351
Db 265 SGASTATNDSSTVSSGA---STATNSESSTTSSGAST-----ATN 302
QY 352 OBEKVSGAGYGL-SGALILGGIGVAVTAALHRKNQPVETTTTTTTTTTTSARTVEN 410
Db 303 SESSTTSSGASTATNDSSTTSSGAGTATNSESSTVSSGISTVTNSESSTPSSGANTATN 362
QY 411 KPANNTPAQGNVDTPGSEDTMESRRSSMASTSTFFDTSSIGTVQNPYADVKT----- 464
Db 363 SESSTTSSGANTAT-NSESSTVSSGASTATNSESSTTSSGVSTATNSESSTTSSGASTAT 421
QY 465 LHDSQVPTNSNTSVQNMGNTDSVYTIQHPROTDTNGARLLGNPAGIQSTVARLAL 524
Db 422 NSDSSTTSEASTATNSESSTVSSGISTVTNSESSTTSSGANTATNATNCGSVTS----- 474
QY 525 SGGLRHDGMGLTGGNSAVNTSNPPARG 553
Db 475 AGSGTAALTGMHTTSHSA-STAVSEAKPG 502

RESULT 14
ABU84417
ID ABU84417 standard; protein; 596 AA.
XX AC ABU84417;
XX DT 02-AUG-2003 (first entry)
XX DE Human secreted/transmembrane protein (PRO) #155.
XX KW Human; secreted and transmembrane protein; PRO; TNF-alpha;
KW tumour necrosis factor alpha; chondrocyte cell; tumour; gene therapy;
KW tissue typing.
XX OS Homo sapiens.
XX PN US2003032112-A1.
XX PD 13-FEB-2003.
XX PF 21-JUN-2002; 2002US-00176756.
XX PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.

PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 09-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088039P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 12-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089633P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.

PR 22-JUN-1998; 98US-0090254P.
 PR 24-JUN-1998; 98US-0090422P.
 PR 24-JUN-1998; 98US-0090433P.
 PR 24-JUN-1998; 98US-0090444P.
 PR 24-JUN-1998; 98US-0090461P.
 PR 24-JUN-1998; 98US-0090535P.
 PR 24-JUN-1998; 98US-0090540P.
 PR 25-JUN-1998; 98US-0090678P.
 PR 25-JUN-1998; 98US-0090678P.
 PR 25-JUN-1998; 98US-0090688P.
 PR 25-JUN-1998; 98US-0090690P.
 PR 25-JUN-1998; 98US-0090694P.
 PR 25-JUN-1998; 98US-0090695P.
 PR 25-JUN-1998; 98US-0090698P.
 PR 26-JUN-1998; 98US-00105413.
 PR 26-JUN-1998; 98US-0090862P.
 PR 26-JUN-1998; 98US-0090863P.
 PR 26-JUN-1998; 98US-0091010P.
 PR 01-JUL-1998; 98US-0091359P.
 PR 01-JUL-1998; 98US-0091544P.
 PR 02-JUL-1998; 98US-0091478P.
 PR 02-JUL-1998; 98US-0091486P.
 PR 02-JUL-1998; 98US-0091626P.
 PR 02-JUL-1998; 98US-0091632P.
 PR 24-JUL-1998; 98US-0094006P.
 PR 04-AUG-1998; 98US-0095282P.
 PR 10-AUG-1998; 98US-0095998P.
 PR 10-AUG-1998; 98US-0096012P.
 PR 17-AUG-1998; 98US-0096757P.
 PR 17-AUG-1998; 98US-0096766P.
 PR 17-AUG-1998; 98US-0096867P.
 PR 17-AUG-1998; 98US-0096891P.
 PR 17-AUG-1998; 98US-0096897P.
 PR 18-AUG-1998; 98US-0096949P.
 PR 18-AUG-1998; 98US-0096959P.
 PR 18-AUG-1998; 98US-0097022P.
 PR 26-AUG-1998; 98US-0097952P.
 PR 26-AUG-1998; 98US-0097954P.
 PR 26-AUG-1998; 98US-0097955P.
 PR 26-AUG-1998; 98US-0097971P.
 PR 26-AUG-1998; 98US-0097974P.
 PR 26-AUG-1998; 98US-0098014P.
 PR 01-SEP-1998; 98US-0098716P.
 PR 01-SEP-1998; 98US-0098723P.
 PR 02-SEP-1998; 98US-0098803P.
 PR 02-SEP-1998; 98US-0098821P.
 PR 02-SEP-1998; 98US-0098843P.
 PR 09-SEP-1998; 98US-0098602P.
 PR 10-SEP-1998; 98US-0099741P.
 PR 10-SEP-1998; 98US-0099754P.
 PR 10-SEP-1998; 98US-0099763P.
 PR 10-SEP-1998; 98US-0099812P.
 PR 15-SEP-1998; 98US-0100388P.
 PR 16-SEP-1998; 98US-0100662P.
 PR 16-SEP-1998; 98US-0100664P.
 PR 16-SEP-1998; 98US-0101751P.
 PR 16-SEP-1998; 98US-0101751P.
 PR 17-SEP-1998; 98US-0100683P.
 PR 17-SEP-1998; 98US-0100684P.
 PR 17-SEP-1998; 98US-0100919P.
 PR 17-SEP-1998; 98US-0100930P.
 PR 18-SEP-1998; 98US-0100849P.
 PR 18-SEP-1998; 98US-0101014P.
 PR 18-SEP-1998; 98US-0101068P.
 PR 23-SEP-1998; 98US-0101471P.
 PR 23-SEP-1998; 98US-0101472P.
 PR 23-SEP-1998; 98US-0101475P.
 PR 23-SEP-1998; 98US-0101477P.
 PR 24-SEP-1998; 98US-0101738P.
 PR 24-SEP-1998; 98US-0101739P.
 PR 24-SEP-1998; 98US-0101743P.
 PR 24-SEP-1998; 98US-0101922P.

PR 25-SEP-1998; 98US-0101786P.
 PR 29-SEP-1998; 98US-0102207P.
 PR 29-SEP-1998; 98US-0102240P.
 PR 29-SEP-1998; 98US-0102330P.
 PR 29-SEP-1998; 98US-0102331P.
 PR 30-SEP-1998; 98US-0102487P.
 PR 30-SEP-1998; 98US-0102570P.
 PR 30-SEP-1998; 98US-0102571P.
 PR 01-OCT-1998; 98US-0102684P.
 PR 01-OCT-1998; 98US-0102687P.
 PR 02-OCT-1998; 98US-0102965P.
 PR 06-OCT-1998; 98US-0103258P.
 PR 06-OCT-1998; 98US-0103449P.
 PR 07-OCT-1998; 98US-00168978.
 Query Match 6.3%; Score 180; DB 6; Length 596;
 Best Local Similarity 19.2%; Pred. No. 0.00044;
 Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;
 QY 4 GNIGHNPNVNSIPPAPPLPSQTDGAGRGQLINSTGLSGRALFTPVNSMADSGDNRA 63
 DB 13 GLLHLEAATNS-----NETSANTGSSVSSG-----ASTATNSG 49
 QY 64 SDVPGLPVNPMRLAASEITLNDGFEVLHDHGPLOTLNRQIGSSV-FRVEIQBDGKHIAVG 122
 DB 50 SSVTSSGVSTATISGSSVTSN-GVSIV-TNSEFHTTSGISTATNSESTASSGISIATN 107
 QY 123 QRNGVETS--VVLSDQEYARLQSIDPECKDFVFTGGRGGAGHAMVTVASDITEARQIL 180
 DB 108 SESSTTSSGASTATNSE---SSTPSSGASTVTVNSGSSVTSSGASTATNSESTVSSRAS 163
 QY 181 ELLEPKGTGSKGAGESKGVGELRESNGAENTTETQTSTSSLSRSPKLWLALGTVAT 240
 DB 164 TATNSESTLSSGASTATN-SDSSTTSSGASTATNSESTTSSG-----ASTAT 211
 QY 241 GLIGLAATGIYOALALTPEDSPPTTDDPAASATETATRDQLTKEAFQNPDPNOKWIDE 300
 DB 212 N-----SESSTVSSRASTATNSESTT---SSGASTATNSESTTNGAGTATNSESTTS 264
 QY 301 LG-----NAIPSGVLKDDVVANIEEOAKAAGEEAKQQAIIENNAQAQKYDQQAQR 351
 DB 265 SGASTATNSDSSSTVSSGA---STATNSESTTSSGAST-----ATN 302
 QY 352 QBELKVSSGAGYGL-SGALILGGGIGVAVTAALHRKNOPVEQTITTTTTTTTTSARTVEN 410
 DB 303 SESSTTSSGASTATNSDSSSTTSSGASTATNSESTVSSGISTVTVNSSESTSSGANTATN 362
 QY 411 KPANNTPAQGNVDTPGSEDVTWESRRSSMASTSTTFDTSIGTVQNPYADVKTST----- 464
 DB 363 SESSTTSSGANTAT-NSESTVSSGASTATNSESTTSSGASTATNSESTTSSGASTAT 421
 QY 465 LHDSQVPTNSNTSVQNMGNNTDSVVYSTIOHPPRDITDNGARLLGNPSAGIQSTYARLAL 524
 DB 422 NSDSSSTTSSSEASTATNSESTVSSGISTVTVNSSESTTSSGANTATNSGSSVTS----- 474
 QY 525 SGGLRHDMGGLTGGSNSAVNTSNPPAPG 553
 DB 475 AGSGTAALTGMHTTSHSA-STAVSEAKPG 502

RESULT 15

ABR66291
 ID ABR66291 standard; protein; 596 AA.

XX ABR66291;

XX 05-AUG-2003 (first entry)

XX Human secreted polypeptide PRO1342, SEQ ID NO:310.

XX Human; PRO; secreted protein; transmembrane protein;
 KW extracellular domain; tumour necrosis factor-alpha; TNF-alpha;
 KW chondrocyte; proliferation; differentiation; cartilage disorder;

bone disorder; arthritis; sports injury; cancer; tumour; diagnosis;
adrenal tumour; lung; colon; breast; prostate; kidney; rectum; cervix;
liver; drug screening; transgenic animal; genetic analysis;
antiarthritic; vulnereary; gene therapy.

Homo sapiens.

US2003027278-A1.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 11-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089050P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096881P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.

```

PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 02-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98US-01019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.

Query Match      6.3%; Score 180; DB 6; Length 596;
Best Local Similarity 19.2%; Pred. No. 0.00044;
Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;

QY 4 GNLCNPNVNNSTPPAPLPQTDGAGRGQLINSTPLGSRALFTPVNSMADSGDNEA 63
DB 13 GULLHLEAATNS-----NETSTANTGSSVSSG-----ASTATNSG 49

QY 64 SDVPGLPVNPMRLAASEITLNDGFEVLHDHGLDPLTLNROIGSSV-PRVETQEDGKHIAVG 122
DB 50 SSVTSSGVATISGSSVTSN-GVSIV-TNSEPHITSSGISTATNSEFSTASSGSIATN 107

QY 123 QRNGVETS--VVLSDQFYARLQSIDPEGKDKFVFTGCGRAGHAMVTVASDITEARQIL 180
DB 108 SSSSTTSSGASTATNSE---SSTPSSGASTVNSGSSVTSNGASTATNSSESTVSSRAS 163

QY 181 ELLEPKGTGCKGAGSKGVGELRENSGAENTTETQTSSTSLRSDPKLWALGTAT 240
DB 164 TATNSESSTLSSGASTATN-SDSSTTSSGASTATNSSESTTSSG-----ASTAT 211

QY 241 GLIGLAATGIVQALATPEPDSPTTDDPAASAATETATRDOLTKEAFQNPONOKNIDE 300
DB 212 N-----SESSTVSSRASTATNSESSTT---SSGASTATNSESRTTNGAGTATNSESSTTS 264

QY 301 LG-----NAIPSGVLKDDVYANIEBQAKAAGEAKQQAIIENNAQAQKYDEQAKR 351
DB 265 SGASTATNSDSSTVSSGA---STATNSESSTTSSGAST-----ATN 302

QY 352 QBELKVSSGAGYGL-SCALILGGIGVAVTAALHRKNQPVETTTTTTTTTTTTTSARTVEN 410

```

```

DB 303 SESSTTSSGASTATNSDSSTTSSGAGTATNSESSTVSSGISTVTSSESTPPSGANTATN 362
QY 411 KPANNTPAQGNVDTPGSDTMESSRRSSMASTSTFFDTSSICTVQNPYADVKTSS----- 464
DB 363 SESSTTSSGANTAT-NSESSTVSSGASTATNSESSTTSSGVSTATNSESSTTSSGASTAT 421
QY 465 LHDSQVPTSNSTSVQNMGNTDSVYVSTIQHPPTTONGARLLGNPNSAGIOSIYARLAL 524
DB 422 NSDSSTTSEASTATNSESSTVSSGISTVTSSESTTSSGASTATNSESSTVSSSVTS----- 474
QY 525 SGGLRHDMMGLTGGSNSAVNTSNPPAPG 553
DB 475 AGSGTAALTGMHTTSHA-STAVSEAKPG 502

```

Search completed: May 13, 2005, 11:44:31
Job time : 172 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)